



INVESTOR IN PEOPLE

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

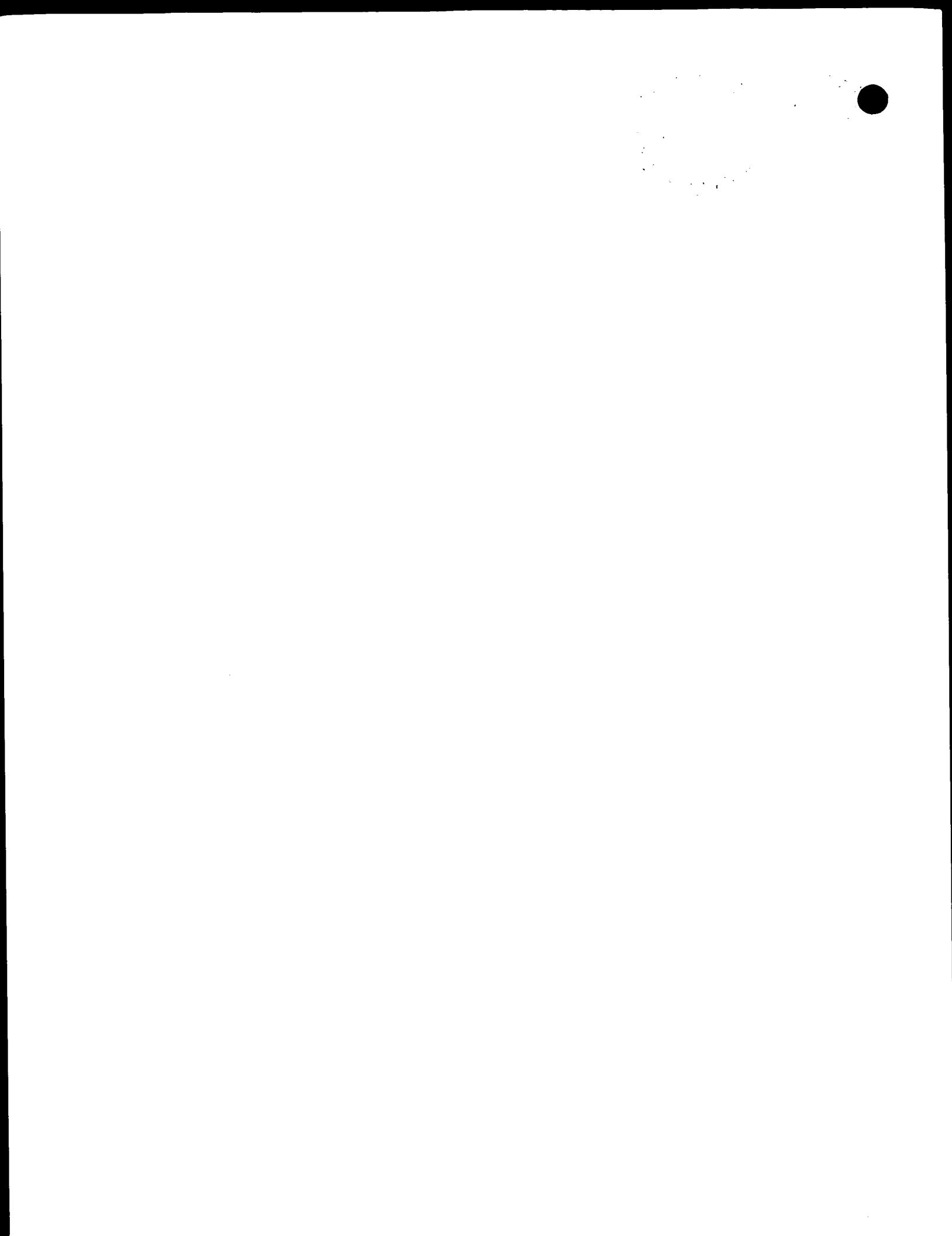
In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed

Dated

Stephen Hordley
18 December 2001

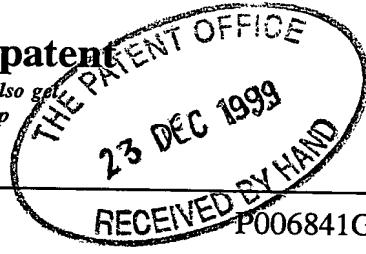


29DEC99 E501744-7 D02246
P01/7700 0.00-9930530.2

Request for a grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

1. Your reference



P006841GBR ATM

The Patent Office

 Cardiff Road
 Newport
 Gwent NP9 1RH

2. Patent application number
(The Patent Office will fill in this part)

9930530.2

23 DEC 1999

3. Full name, address and postcode of the or of each applicant
(underline all surnames)

Medical Research Council
 20 Park Crescent
 London
 W1N 4AL
 United Kingdom

Patents ADP number (if you know it) 596007001

If the applicant is a corporate body, give the country/state of its incorporation United Kingdom

4. Title of the invention

Oligomeric Chaperone Proteins

5. Name of your agent (if you have one)

D YOUNG & CO

"Address for service" in the United Kingdom to which all correspondence should be sent
(including the postcode)

21 NEW FETTER LANE
 LONDON
 EC4A 1DA

Patents ADP number (if you have one)

59006 ✓

6. If you are declaring priority from one or more earlier patent applications, give the country and date of filing of the or each of these earlier applications and (if you know it) the or each application number

Country

Priority application number
(if you know it)

Date of filing
(day/month/year)

7. If this application is divided or otherwise derived from an earlier UK application, give the number and filing date of the earlier application

Number of earlier application

Date of filing
(day/month/year)

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? (Answer 'Yes' if:
a) any applicant named in part 3 is not an inventor, or
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body.
See note (d))

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form 0

Description 53

Claims(s) 4

Abstract 1

Drawing(s) 910

20

X 10

10. If you are also filing any of the following, state how many against each item.

Priority documents

Translations of priority documents

Statement of inventorship and right to grant of a patent (Patents Form 7/77)

Request for preliminary examination and search (Patents Form 9/77)

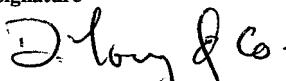
Request for substantive examination (Patents Form 10/77)

Any other documents (please specify)

11.

I/We request the grant of a patent on the basis of this application.

Signature



Date

23 12 99

D YOUNG & CO
Agents for the Applicants

12. Name and daytime telephone number of the person to contact in the United Kingdom

Dr A Maschio

023 8063 4816

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

- a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 01645 500505
- b) Write your answers in capital letters using black ink or you may type them
- c) If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.
- d) If you answered 'Yes' Patents Form 7/77 will need to be filed.
- e) Once you have filled in the form you must remember to sign and date it.
- f) For details of the fee and ways to pay please contact the Patent Office.

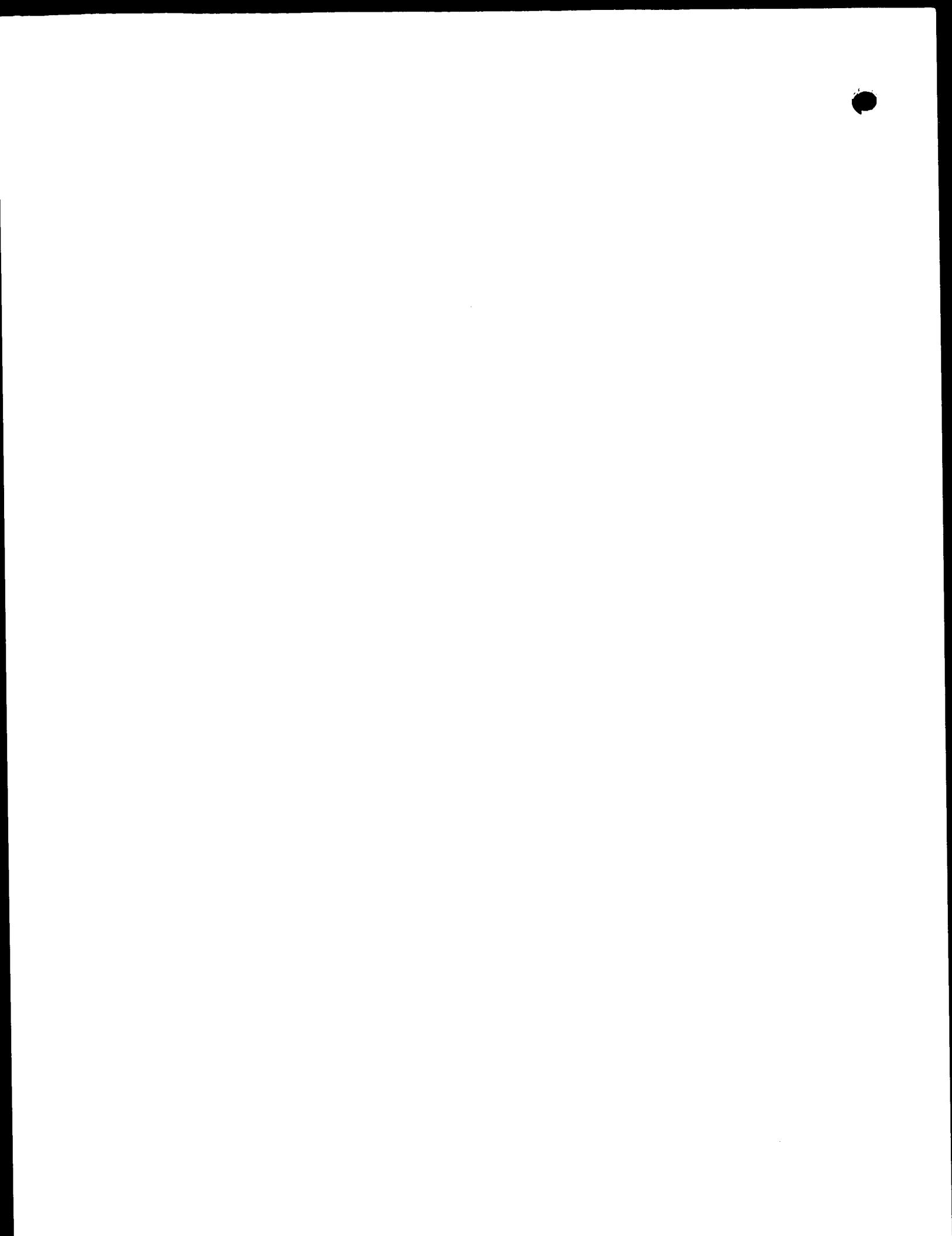
Oligomeric Chaperone Proteins

The present invention relates to oligomers of chaperone or minichaperone proteins. In particular, the invention relates to polypeptides constructed by oligomerisation of 5 minichaperone monomers to form a heptameric ring structure.

Proteins, especially catalytic proteins (enzymes) and proteins which have biological activities, are dependent on tertiary structure for most or all of their functional attributes. Tertiary structure, which is defined by the three-dimensional arrangement of the protein, 10 is dependent on the folding of the primary polypeptide sequence in three dimensions. The tertiary structure is stabilised by the interaction between parts of the primary sequence in the folded state, such as the formation of disulphide bonds, and energetic considerations deriving from juxtaposition of particular chemical entities in the three-dimensional arrangement.

15 It is known that the tertiary structure of proteins, particularly proteins which are stored for any length of time, can degrade, leading to suboptimal activity. This can be due to a variety of factors, including aggregation and the formation of improper intra- and inter-molecular interactions. Moreover, it is also known that proteins produced by recombinant 20 DNA technology are frequently misfolded, especially if produced in bacterial expression systems. The strongly reducing environment present in bacterial cytoplasm impedes the formation of correct disulphide bonds, thus obstructing the folding process.

Many proteins require the assistance of molecular chaperones in order to be fold *in vivo* 25 or to be refolded *in vitro* in high yields. Molecular chaperones are proteins, which are often large and require an energy source such as ATP to function. A key molecular chaperone in *Escherichia coli* is GroEL, which consists of 14 subunits each of some 57.5 kD molecular mass arranged in two seven membered rings. There is a large cavity in the GroEL ring system, and it is widely believed that the cavity is required for successful 30 protein folding activity. For optimal activity, a co-chaperone, GroES, is required which consists of a seven membered ring of 10 kD subunits. The activity of the GroEL/GroES complex requires energy source ATP.



Some proteins are monomers, consisting of a single subunit. Many proteins are oligomeric, consisting of more than one subunit. Sometimes the subunits are identical, sometimes there are different types of subunits. Frequently, the subunits are linked non-covalently. Sometimes, the subunits are attached covalently, with a stretch of polypeptide 5 linking the C-terminus of one domain to the N-terminus of another.

Allosteric proteins are a special class of oligomeric proteins, which alternate between two or more different three-dimensional structures on the binding of ligands and substrates. Allosteric proteins are often involved in control processes in biology or where mechanical 10 and physico-chemical energies are interconverted.

GroEL is an allosteric protein. The role of ATP is to trigger this allosteric change, causing GroEL to convert from a state that binds denatured proteins tightly to one that binds denatured proteins weakly. The co-chaperone, GroES, aids in this process by favouring 15 the weak-binding state. It may also act as a cap, sealing off the cavity of GroEL. Further, its binding to GroEL is likely directly to compete with the binding of denatured substrates. The net result is that the binding of GroES and ATP to GroEL which has a substrate bound in its denatured form is to release the denatured substrate either into the cavity or into solution where it can refold.

20

Minichaperones have been described in detail elsewhere (see International patent application WO99/05163, the disclosure of which is incorporated herein by reference). Minichaperone polypeptides possess chaperoning activity when in monomeric form and do not require energy in the form of ATP. Defined fragments of the apical domain of 25 GroEL of approximately 143-186 amino acid residues in length have molecular chaperone activity towards proteins either in solution under monomeric conditions or when monodisperse and attached to a support.

30 The activity of minichaperones, although sufficient for many purposes, is inferior to that of intact GroEL. There is thus a need for a more active form of minichaperone, which nevertheless retains independence from energy requirements.

The activity of minichaperones, although sufficient for many purposes, is inferior to that of intact GroEL. It is postulated that this could be due to the inability of GroES to oligomerise. There is thus a widespread requirement for a system which would allow the oligomerisation of polypeptides to form functional protein oligomers which have oligomerisation, said monomer comprising a polypeptide sequence which potentiates oligomerisation, said monomer comprising a polypeptide sequence which potentiates protein folding inserted into the sequence of a subunit of an oligomerisable protein according to the present invention, there is provided a polypeptide monomer capable of activating those of recombinant monomeric polypeptides.

It has been observed that the oligomerisation of polypeptide sequences which potentiate protein folding polypeptides allows their spatial juxtaposition, which potentiates their activity.

A "polypeptide sequence which potentiates protein folding" may any one of a number of polyptides which are capable of facilitating the correct folding or refolding of proteins or polypeptides *in vivo* and/or *in vitro*. For example, the polypeptide sequence which may be a minichaperone polypeptide or a foldase. Potentiates protein folding may be a minichaperone which foldase is selected from the group consisting of *E. coli* Dsba and mammalian oxidoreductases and peptidyl prolyl isomerases. Preferably, the thiol/disulphide oxidoreductases and peptidyl prolyl isomerases. Preferably, the thiol/disulphide oxidoreductase is selected from the group consisting of thiol/disulphide Advantagously, the foldase is selected from the group consisting of thiol/disulphide oxidoreductases and peptidyl prolyl isomerases. Preferably, the thiol/disulphide oxidoreductase is selected from the group consisting of *E. coli* Dsba and mammalian PD1, or a derivative thereof. Preferably, the peptidyl prolyl isomerase is a cyclophilin.

In a preferred aspect, the polypeptide sequence which potentiates protein folding may be a protease protease. Whereas molecular chaperones do not provide steric information a protease protease. Whereas molecular chaperones do not provide steric information necessary for newly synthesised proteins to fold correctly, but minimise the risk that partially folded protein chains will aggregate with one another, protease proteases, which are intermolecular chaperones, provide steric information indispensable for the rest of the protein to fold correctly.

20 A "polypeptide sequence which potentiates protein folding" may any one of a number of polyptides which are capable of facilitating the correct folding or refolding of proteins or polypeptides *in vivo* and/or *in vitro*. For example, the polypeptide sequence which which foldase is selected from the group consisting of a minichaperone polypeptide or a foldase. Potentiates protein folding may be a minichaperone which foldase is selected from the group consisting of a minichaperone polypeptide or a foldase. Preferably, the thiol/disulphide oxidoreductase is selected from the group consisting of thiol/disulphide oxidoreductases and peptidyl prolyl isomerases. Preferably, the thiol/disulphide oxidoreductase is selected from the group consisting of *E. coli* Dsba and mammalian PD1, or a derivative thereof. Preferably, the peptidyl prolyl isomerase is a cyclophilin.

25 A "polypeptide sequence which potentiates protein folding" may any one of a number of polyptides which are capable of facilitating the correct folding or refolding of proteins or polypeptides *in vivo* and/or *in vitro*. For example, the polypeptide sequence which which foldase is selected from the group consisting of a minichaperone polypeptide or a foldase. Potentiates protein folding may be a minichaperone which foldase is selected from the group consisting of a minichaperone polypeptide or a foldase. Preferably, the thiol/disulphide oxidoreductase is selected from the group consisting of thiol/disulphide oxidoreductases and peptidyl prolyl isomerases. Preferably, the thiol/disulphide oxidoreductase is selected from the group consisting of *E. coli* Dsba and mammalian PD1, or a derivative thereof. Preferably, the peptidyl prolyl isomerase is a cyclophilin.

30 A "polypeptide sequence which potentiates protein folding" may any one of a number of polyptides which are capable of facilitating the correct folding or refolding of proteins or polypeptides *in vivo* and/or *in vitro*. For example, the polypeptide sequence which which foldase is selected from the group consisting of a minichaperone polypeptide or a foldase. Potentiates protein folding may be a minichaperone which foldase is selected from the group consisting of a minichaperone polypeptide or a foldase. Preferably, the thiol/disulphide oxidoreductase is selected from the group consisting of thiol/disulphide oxidoreductases and peptidyl prolyl isomerases. Preferably, the thiol/disulphide oxidoreductase is selected from the group consisting of *E. coli* Dsba and mammalian PD1, or a derivative thereof. Preferably, the peptidyl prolyl isomerase is a cyclophilin.

35 The activity of minichaperones, although sufficient for many purposes, is inferior to that of intact GroEL. It is postulated that this could be due to the inability of GroES to oligomerise. There is thus a widespread requirement for a system which would allow the oligomerisation of polypeptides to form functional protein oligomers which have oligomerisation, said monomer comprising a polypeptide sequence which potentiates oligomerisation, said monomer comprising a polypeptide sequence which potentiates protein folding inserted into the sequence of a subunit of an oligomerisable protein according to the present invention, there is provided a polypeptide monomer capable of activating those of recombinant monomeric polypeptides.

Summary of the Invention

A "minichaperone polypeptide", as referred to herein, refers to a minichaperone polypeptide as described in WO99/05163, the disclosure of which is incorporated herein by reference. The minichaperone polypeptide preferably comprises fragments of a molecular chaperone, preferably fragments of any hsp-60 chaperone, and may be selected 5 from the group consisting of mammalian hsp-60 and GroEL, or a derivative thereof.

In the case that the fragment is a fragment of GroEL, it advantageously does not have an Alanine residue at position 262 and/or an Isoleucine residue at position 267 of the sequence of intact GroEL as defined at GenBank Accession No. P06159. Preferably, it 10 has a Leucine residue at position 262 and/or a Methionine residue at position 267 of the sequence of intact GroEL. The invention therefore encompasses the use of a fragment of GroEL comprising a Leucine residue at position 262 and/or a Methionine residue at position 267 of the sequence of intact GroEL for potentiating the folding of a polypeptide.

15 In a preferred embodiment, the minichaperone comprises a region which is homologous to at least one of fragments 191-376, 191-345 and 191-335 of the sequence of intact GroEL.

A protein scaffold is a protein, or part thereof, whose function is to determine the 20 structure of the protein itself, or of a group of associated proteins or other molecules. Scaffolds therefore have a defined three-dimensional structure when assembled, and have the capacity to support molecules or polypeptide domains in or on the said structure. Advantageously, a scaffold has the ability to assume a variety of viable geometries, in relation to the three-dimensional structure of the scaffold and/or the insertion site of the 25 polypeptides.

Preferably, the scaffold according to the invention is a chaperonin cpn10/Hsp10 scaffold. Cpn10 is a widespread component of the cpn60/cpn10 chaperonin system. Examples of cpn10 include bacterial GroES and bacteriophage T4 Gp31. Further members of the 30 cpn10 family will be known to those skilled in the art.

The invention moreover comprises the use of derivatives of naturally-occurring scaffolds. Derivatives of scaffolds (including scaffolds of the cpn10 and 60 families) comprise

preferably 44 to 62, advantageously 56 to 57 of *E. coli* GroES. 30
to 66, and preferably 59 and 61 of bacteriophage T4 Gp31, or between positions 43 to 63, family peptides. This position is located between positions 54 and 67, advantageously 55 terminus thereof, or in positions which are equivalent to the root of hairpin of cpn10 sequence which potentiates protein folding may moreover be incorporated at the N or C 25 where the protein scaffold subunit is a cpn10 family polypeptide, the polypeptide family peptide.

cpn10 family polypeptide. 25
Potentiates protein folding is inserted by replacing all or part of the mobile loop of a 22 to 45, advantageously 23 to 44. Advantageously, the polypeptide sequence which members of the cpn10 family. The mobile loop of T4 Gp31 is located between residues acids 16 to 33, of the sequence of *E. coli* GroES, and equivalent positions on other 20 mobile loop is positioned between amino acids 15 and 34, preferably between amino It is known that cpn10 subunits possess a "mobile loop" within their structure. The scaffold subunit, for example by replacing one or more amino acids thereof.

protein scaffold subunit. Thus, the polypeptide is included with the sequence of the 15 C termini of the polypeptide monomer are formed by the sequence of the oligomeric stable into the sequence of the oligomeric stable protein scaffold subunit such that both the N and advantageously, the polypeptide sequence which potentiates protein folding is inserted 10 scaffold subunit, the scaffold is a seven-membered ring.

Advantageously, therefore, the scaffold is a seven-membered ring. 10
members comprises seven subunits, in the shape of a seven-membered ring or annulus. Between 5 and 15 subunits, and ideally about 10 subunits. The scaffold of cpn10 family subunits. Preferably, the scaffold comprises between 2 and 20 subunits, advantageously present invention, the scaffold may have any shape and may comprise any number of 15 protein scaffold subunits assemble to form a protein scaffold. In the context of the to the maintenance of the "oligomerization" property described herein.

members of the Cpnl0 or 60 families and/or circular permuted protein scaffolds, subject (especially replacement of Cys residues in Gp31), hybrids formed by fusion of different mutants thereof, which may contain amino acid deletions, additions or substitutions 20

Advantageously, the polypeptide sequence which potentiates protein folding may be inserted at an N or C terminus of a scaffold subunit in association with circular permutation of the subunit itself. Circular permutation is described in Graf and Schachman, PNAS(USA) 1996, 93:11591. Essentially, the polypeptide is circularised by fusion of the existing N and C termini, and cleavage of the polypeptide chain elsewhere to create novel N and C termini. In a preferred embodiment of the invention, the polypeptide may be included at the N and/or C terminus formed after circular permutation. The site of formation of the novel termini may be selected according to the features desired, and may include the mobile loop and/or the roof β hairpin.

10

Advantageously, polypeptide sequences which potentiate protein folding, which may be the same or different, may be inserted at more than one of the positions above-identified in the protein scaffold subunit. Thus, each subunit may comprise two or more polypeptides, which are displayed on the scaffold when this is assembled.

15

Polypeptide sequences which potentiate protein folding may be displayed on a scaffold subunit in free or constrained form, depending on the degree of freedom provided by the site of insertion into the scaffold sequence. For example, varying the length of the sequences flanking the mobile loop in the scaffold will modulate the degree of constraint 20 of any polypeptide inserted therein.

In a second aspect, the invention relates to a polypeptide oligomer comprising two or more monomers according to the first aspect of the invention. The oligomer may be configured as a heterooligomer, comprising two or more different polypeptide sequences 25 which potentiate protein folding inserted into the scaffold, or as a homooligomer, in which the polypeptides inserted into the scaffold are the same.

If the oligomer according to the invention is a heterooligomer, it may be configured such that the polypeptides juxtaposed thereon have complementary biological activities. For 30 example, one or more minichaperones and foldases are advantageously displayed on the same scaffold, enabling them to act in concert.

The monomers which constitute the oligomer may be covalently crosslinked to each other. Cross linking may be performed by recombinant approaches, such that the monomers are expressed *ab initio* as an oligomer; alternatively, cross-linking may be performed at Cys residues in the scaffold. For example, unique Cys residues inserted between positions 56 and 57 of the GroES scaffold, or equivalent positions on other members of the cpn10 family, may be used to cross-link scaffold subunits.

In a third aspect, the present invention relates to a method for preparing a polypeptide monomer capable of oligomerisation according to the first aspect of the invention, comprising the steps of inserting a nucleic acid sequence encoding a polypeptide sequence which potentiates protein folding into a nucleic acid sequence encoding a subunit of an oligomerisable protein scaffold, incorporating the resulting nucleic acid into an expression vector, and expressing the nucleic acid to produce the polypeptide monomer.

Oligomericised protease precursors form polyvalent steric chaperones. The invention therefore provides polyvalent protease polypeptides, optionally incorporating one or more minichaperones and/or foldases, which are useful in refolding polypeptides *in vitro* and *in vivo*. Oligomericised proteases may be used to potentiate the folding of, *inter alia*, proteases. Moreover, polyvalent protease polypeptides may be used to alter the folding patterns of polypeptides, thus permanently altering their activities.

The invention moreover relates to a method for producing a polypeptide oligomer according to the second aspect of the invention, comprising allowing the polypeptide monomers produced as above to associate into an oligomer. Preferably, the monomers are cross-linked to form the oligomer.

According to a fourth aspect of the present invention, there is provided a method for potentiating the folding of a polypeptide comprising contacting the polypeptide with a multimeric minichaperone polypeptide as described above.

The polypeptide which is folded by the method of the invention is preferably an unfolded or misfolded polypeptide, and advantageously comprises a disulphide.

The invention moreover concerns a method as described above wherein the oligomer is immobilised onto a solid phase support, which may be agarose. Accordingly, the invention also provides a solid phase support having immobilised thereon an oligomer 5 according to the present invention, and a column packed at least in part with such a solid phase support.

The oligomer according to the invention may be combined with different oligomers, or with independent molecules capable of potentiating protein folding. For example, the 10 oligomers according to the invention may be used, in solution or in immobilised form, in combination with foldases, chaperones or other enzymes.

Brief Description of the Drawings

15 **Figure 1.** (a) Three-dimensional structure of Gp31 of bacteriophage T4 solved at 2.3 Å. Positions mentioned in the text are indicated (residues numbered as in van der Vies, S., Gatenby, A. & Georgopoulos, C. (1994) *Nature* 368, 654-656). (b) Three-dimensional structure of minichaperone GroEL(191-376) solved at 1.7 Å. The distance between residues 25 and 43 of Gp31 is around 12 Å; the distance between residues 191 and 376 of 20 GroEL is around 9 Å. Positions mentioned in the text are indicated (residues numbered as in Hemmingsen, S. M., Woolford, C., van, d. V. S., Tilly, K., Dennis, D. T., Georgopoulos, C. P., Hendrix, R. W. & Ellis, R. J. (1988) *Nature* 333, 330-334.). Secondary structure representations are drawn with MolScript (Kraulis, P. (1991) *J. Appl. Crystallogr.* 24, 946-950).

25

Figure 2. Schematic representation of Gp31 proteins in the vectors used in this study. The presence of the Gp31 mobile loop (residues 23 to 44) and/or minichaperone GroEL (residues 191 to 376) are indicated by boxes. The nucleotide sequence of the Gp31 mobile loop and relevant restriction sites are shown. The names of the corresponding 30 vector are listed in the left margin.

Figure 3. (a) Molecular weight determination by analytical gel filtration chromatography. Wild-type proteins Gp31 ($M_r \approx 7 \times 12$ kDa) and GroEL(191-376) ($M_r \approx 22$ kDa) and,

GP31A loop and GP31A::GroEL(191-376) (MC₇) were run on a Superdex™ 200 (HR 10/30) column (Pharmacia Biotech), calibrated with molecular weight standards (solid-line and circles). GP31A loop and MC₇ eluted at volumes corresponding to molecular weights of \approx 145.6 and \approx 215 kDa, respectively. (b) Molecular weight determination of MC₇ by equilibrium analytical ultracentrifugation. The apparent molecular weight of MC₇ is \approx 215 kDa.

Figure 4. Characterization of MC₇ by CD spectroscopy. (a) Far UV-CD spectrum at 25 °C. (b) Thermal denaturation followed at 222 nm at a heating rate of 1 °C·min⁻¹.

Figure 5. (a) Binding specificity of MC₇ to GroES determined by ELISA. (b) Inhibition of MC₇ binding to heptameric co-chaperonin GroES by varying concentrations of synthetic peptide corresponding to residues 16 to 32 of GroES mobile loop determined by competition ELISA.

Figure 7. *In vitro* refolding of heat- and dithiothreitol-denatured mtMDH. (a) Protection of aggregation at 47 °C followed by light scattering at 550 nm. (b) Time-dependent reactivation of mtMDH at 25 °C. (c) Yields of mtMDH reactivation.

Figure 8 and Figure 9 show the possible insertion sites for polypeptides in bacteophagelike Figure 10 illustrates the potential attachment sites for polypeptides to a circular scaffold, in this case GP31.

25

TA GP31.

Figure 8 and Figure 9 show the possible insertion sites for polypeptides in bacteophagelike

15

10

5

Detailed Description of the Invention**Definitions**

5 *Oligomerisable scaffold.* An oligomerisable scaffold, as referred to herein, is a polypeptide which is capable of oligomerising to form a scaffold and to which a polypeptide may be fused, preferably covalently, without abolishing the oligomerisation capabilities. Thus, it provides a “scaffold” using which polypeptides may be arranged into multimers in accordance with the present invention. Optionally, parts of the wild-
10 type polypeptide from which the scaffold is derived may be removed, for example by replacement with the polypeptide which is to be presented on the scaffold.

15 *Monomer.* Monomers according to the present invention are polypeptides which possess the potential to oligomerise. This is brought about by the incorporation, in the polypeptide, of an oligomerisable scaffold subunit which will oligomerise with further scaffold subunits if combined therewith.

20 *Oligomer.* As used herein, “oligomer” is synonymous with “polymer” or “multimer” and is used to indicate that the object in question is not monomeric. Thus, oligomeric polypeptides according to the invention comprise at least two monomeric units joined together covalently or non-covalently. The number of monomeric units employed will depend on the intended use of the oligomer, and may be between 2 and 20 or more. Advantageously, it is between 5 and 10, and preferably about 7.

25 *Polypeptide.* As used herein, a polypeptide is a molecule comprising at least one peptide bond linking two amino acids. This term is synonymous with “protein” and “peptide”, both of which are used in the art to describe such molecules. A polypeptide may comprise other, non-amino acid components. The polypeptide the folding of which is potentiated by the method of the invention may be any polypeptide. Preferably, however,
30 it is an unfolded or misfolded polypeptide which is in need of folding. Alternatively, however, it may be a folded polypeptide which is to be maintained in a folded state (see below).

amino acids in length and most preferably about 150 amino acids in length. Fragments of 30 fragments are between 50 and 200 amino acids in length, preferably between 100 and 200 solution. Preferred fragments are described below. Advantageously, chaperone activity. Advantageously, a fragment of a chaperonin molecule remains chaperonin entire native molecular chaperone molecule which nevertheless retains chaperonin 35 *Fragment*. When applied to chaperone molecules, a fragment is anything other than the

25 which may be added to a solution comprising a polypeptide whose folding is to be a column. Alternatively, the solid support may be in the form of beads or another matrix and/or the thiol/disulphide oxidoreductase immobilised on a solid support, for example on solid supports. In a preferred aspect, the contact occurs with the minichaperone oligomer solution, *in vitro* or *in vivo*, with one or more components of the reaction immobilised on 20 with the polypeptides whose folding is to be potentiated. This contact may occur in free *Contacting*. The reagents used in the method of the invention require physical contact 25 eventuates are covered by the reference to "potentiation", the folding of the polypeptide. activity of an already substantially correctly folded polypeptide. These, and other, the folded/unfolded equilibrium so as to favour the folded state. This prevents loss of method of the invention serves to maintain the folded state of the polypeptide by affecting state, that is all or most of it is folded correctly or nearly correctly. In this case, the 30 In this case, its correct folding is potentiated by the method of the invention. A second situation is one in which the polypeptide is substantially already in its correctly folded situation, and many toxins and proteases.

5 Examples of polypeptides include those used for medical or biotechnological use, such as interleukins, interferons, antibodies and their fragments, insulin, transforming growth factor, and many toxins and proteases.

50 Preferably, the polypeptide which is folded according to the invention contains at least one disulphide. Such polypeptides may be referred to herein as *disulphide-containing polypeptides*.

chaperone molecules which remain monomeric in solution and possess a chaperoning activity which is not energy-dependent are referred to as minichaperones.

5 *Unfolded*. As used herein, a polypeptide may be unfolded when at least part of it has not yet acquired a correct or desired secondary or tertiary structure. A polypeptide is *misfolded* when it has acquired an at least partially incorrect or undesired secondary or tertiary structure.

10 *Immobilised, immobilising*. Permanently attached, covalently or otherwise. In a preferred aspect of the present invention, the term "immobilise", and grammatical variations thereof, refer to the attachment of molecular chaperones or, preferably, foldase polypeptides to a solid phase support using a method as described in WO99/05163.

15 *Solid (phase) support*. Reagents used in the invention may be immobilised onto solid phase supports. This means that they are permanently attached to an entity which remains in a different (solid) phase from reagents which are in solution. For example, the solid phase could be in the form of beads, a "polypeptide chip", a resin, a matrix, a gel, the material forming the walls of a vessel or the like. Matrices, and in particular gels, such as agarose gels, may conveniently be packed into columns. A particular advantage 20 of solid phase immobilisation is that the reagents may be removed from contact with the polypeptide(s) with facility.

25 *Foldase*. In general terms, a foldase is an enzyme which participates in the potentiation of protein folding through its enzymatic activity to catalyse the rearrangement or isomerisation of bonds in the folding polypeptide. They are thus distinct from a molecular chaperone, which bind to polypeptides in unstable or non-native structural states and potentiate correct folding without enzymatic catalysis of bond rearrangement. Many 30 classes of foldase are known, and they are common to animals, plants and bacteria. They include peptidyl prolyl isomerases and thiol/disulphide oxidoreductases. The invention comprises the use of all foldases which are capable of potentiating protein folding through covalent bond rearrangement.

Peptidyl-prolyl isomerase. Peptidyl-prolyl isomerasers are known enzymes widely present in a variety of cells. Examples include cyclophilin (see, for example, Bergsma et al. (1991) *J. Biol. Chem.* 266:23204-23214), parvulin, SraA (Rouviere and Gross, (1996) *Genes Dev.* 10:3170-3182) and FK506 binding proteins FKBP51 and FKBP52. PPI is responsible for the *cis-trans* isomerization of peptidyl-prolyl bonds in polypeptides, thus potentially controlling correct folding. The invention includes any polypeptide having PPI activity.

This thiol/disulphide oxidoreductase. As the name implies, thiol/disulphide oxidoreductases catalyse the formation of disulphide bonds and can thus dictate the folding rate of disulphide-containing polypeptides. The invention accordingly comprises the use of any polypeptide possessing such an activity. This includes chaperone polypeptides, or disulphide-containing polypeptides. The invention accordingly comprises the use of any fragments thereof, which may possess PDI activity (Wang & Tsou, (1998) FEBS lett. 425:382-384). In Eukaryotes, thiol/disulphide oxidoreductases are generally referred to as PDIs (protein disulphide isomerases). PDI interacts directly with newly synthesised secretory proteins and is required for the folding of nascent polypeptides in the endoplasmic reticulum (ER) of eukaryotic cells. Enzymes found in the ER with PDI activity include mammalian PDI (Edman et al., 1985, Nature 317:267, yeast PDI (Mizunaga et al. 1990, J. Biocchem. 108:848), mammalian ERp59 (Mazzarella et al., 1990, J. Biocchem. 265:1094), mammalian prolyl-4-hydroxylase (Pihlajaniemi et al., 1987, EMBO J. 6: 643) yeast GSBP (Lamantia et al., 1991, Proc. Natl. Acad. Sci. USA, 88:4453) and mammalian T3BP (Yamuchi et al., 1987, Biocchem. Biophys. Res. Commun. 146:1485), A. niger PdiA (Ngiama et al., (1997) Curr. Genet. 31:133-138) and yeast EUG1 (Tachibana et al., 1992, Mol. Cell Biol. 12, 4601). In prokaryotes, equivalent proteins exist, such as the Dsba protein of *E. coli*. Other peptides with similar activity include, for example, p52 from *T. cruzi* (Moutiez et al., (1997) Biocchem. J. 322:43-48). These polypeptides, and other functionally equivalent polypeptides, are included with the scope of the present invention, as are derivatives of the polypeptides which share the relevant activity (see below). Preferably, the thiol/disulphide oxidoreductase according to the invention is selected from the group consisting of mammalian PDI or *E. coli* Dsba.

Moreover, as used herein, the term "a foldase" includes one or more foldases. In general, in the present specification the use of the singular does not preclude the presence of a plurality of the entities referred to, unless the context specifically requires otherwise.

This includes chaperone polypeptides, or fragments thereof, which may possess PPI activity (Wang & Tsou, (1998) FEBS lett. 425:382-384).

Molecular Chaperone. Chaperones, or chaperonins, are polypeptides which potentiate protein folding by non-enzymatic means, in that they do not catalyse the chemical modification of any structures in folding polypeptides, by potentiate the correct folding of polypeptides by facilitating correct structural alignment thereof. Molecular chaperones are well known in the art, several families thereof being characterised. The invention is applicable to any molecular chaperone molecule, which term includes, for example, the molecular chaperones selected from the following non-exhaustive group:

p90 Calnexin	Salopek <i>et al.</i> , J. Investig Dermatol Symp Proc (1996) 1:195
HSP family	Walsh <i>et al.</i> , Cell Mol. Life Sci. (1997) 53:198
HSP 70 family	Rokutan <i>et al.</i> , J. Med. Invest. (1998) 44:137
DNA K	Rudiger <i>et al.</i> , Nat. Struct. Biol. (1997) 4:342
DNAJ	Cheetham <i>et al.</i> , Cell Stress Chaperones (1998) 3:28
HSP 60 family; GroEL	Richardson <i>et al.</i> , Trends Biochem. (1998) 23:138
ER-associated chaperones	Kim <i>et al.</i> , Endocr Rev (1998) 19:173
HSP 90	Smith, Biol. Chem. (1988) 379:283
Hsc 70	Hohfeld, Biol. Chem. (1988) 379:269
sHsps; SecA; SecB	Beissinger <i>et al.</i> , Biol. Chem. (1988) 379:245
Trigger factor	Wang <i>et al.</i> , FEBS Lett. (1998) 425:382
zebrafish hsp 47, 70 and 90	Krone <i>et al.</i> , Biochem. Cell Biol. (1997) 75:487
HSP 47	Nagata, Matrix Biol. (1998) 16:379
GRP 94	Nicchitta <i>et al.</i> , Curr. Opin. Immunol. (1998) 10:103
Cpn 10	Cavanagh, Rev. Reprod. (1996) 1:28
BiP	Sommer <i>et al.</i> , FASEB J. (1997) 11:1227
GRP 78	Brostrom <i>et al.</i> , Prog. Nucl. Acid. res. Mol. Biol. (1998) 58:79
C1p, FtsH	Suzuki <i>et al.</i> , Trends Biochem. Sci. (1997) 22:118
Ig invariant chain	Weenink <i>et al.</i> Immunol. Cell biol. (1997) 75:69
mitochondrial hsp 70	Horst <i>et al.</i> , BBA (1997) 1318:71

Two major families of protein folding chaperones which have been identified, the heat shock protein 60 (hsp60) class and the heat shock protein 70 (hsp70) class, are especially preferred for use herein. Chaperones of the hsp-60 class are structurally distinct from chaperones of the hsp-70 class. In particular, hsp-60 chaperones appear to form a stable scaffold of two heptamer rings stacked one atop another which interacts with partially folded elements of secondary structure. On the other hand, hsp-70 chaperones are monomers of dimers and appear to interact with short extended regions of a polypeptide.

10 Hsp70 chaperones are well conserved in sequence and function. Analogs of hsp-70 include the eukaryotic hsp70 homologue originally identified as the IgG heavy chain endoplasmic reticulum (ER). Bip is located in all eukaryotic cells within the lumen of the ER. Esccherichia coli shares about 50% sequence homology with an hsp70 KAR2 chaperone in yeast (Rose *et al.* 1989 Cell 57:1211-1221). Moreover, the presence of mouse Bip in yeast can functionally replace a lost yeast KAR2 gene (Nomigton *et al.* 19: 1223-1236).

20 Hsp-60 chaperones are universally conserved (Zeilstra-Ryalls *et al.*, 1991) Ann. Rev. Microbiol. 45:301-325) and include hsp-60 homologues from large number of species, including man. They include, for example, the *E. coli* GroEL polypeptide, *Escherichia* *sennei* GroEL (Zhang *et al.*, 1997) FEMS Immunol. Med. Microbiol. 18:39-46); *Trichomonas vaginalis* hsp-60 (Bozner *et al.*, 1997) J. Parasitol. 83:224-229; rat hsp-60

(Venner *et al.*, (1990) NAR 18:5309; and yeast hsp-60 (Johnson *et al.*, (1989) Gene 84:295-302.

5 In a preferred aspect, the present invention relates to fragments of polypeptides of the hsp-60 family. These proteins being universally conserved, any member of the family may be used; however, in a particularly advantageous embodiment, fragments of GroEL, such as *E. coli* GroEL, are employed. It has also found that agarose-immobilised calmodulin does have a chaperoning activity, presumably because of its exposed hydrophobic groups.

10

The sequence of GroEL is available in the art and from academic databases (see GenBank Accession No. P06159); however, GroEL fragments which conform to the database sequence are inoperative. Specifically, the database contains a sequence in which positions 262 and 267 are occupied by Alanine and Isoleucine respectively. Fragments 15 incorporating one or both of these residues at these positions are inoperative and unable to potentiate the folding of polypeptides. The invention, instead, relates to a GroEL polypeptide in which at least one of positions 262 and 267 is occupied by Leucine and Methionine respectively.

20

Derivative. The present invention relates to derivatives of molecular chaperones, peptidyl-prolyl isomerases and thiol/disulphide oxidoreductases. In a preferred aspect, therefore, the terms "molecular chaperone", "peptidyl-prolyl isomerase" and "thiol-disulphide oxidoreductase" include derivatives thereof which retain the stated activity. The derivatives provided by the present invention include splice variants encoded by 25 mRNA generated by alternative splicing of a primary transcript, amino acid mutants, glycosylation variants and other covalent derivatives of molecular chaperones or foldases which retain the functional properties of molecular chaperones, peptidyl-prolyl isomerases and/or thiol/disulphide oxidoreductases. Exemplary derivatives include molecules which are covalently modified by substitution, chemical, enzymatic, or other 30 appropriate means with a moiety other than a naturally occurring amino acid. Such a moiety may be a detectable moiety such as an enzyme or a radioisotope. Further included are naturally occurring variants of molecular chaperones or foldases found within a particular species, whether mammalian, other vertebrate, yeast, prokaryotic or otherwise.

10 In a preferred embodiment, the scaffold polypeptide is based on members of the cpn10/Hsp10 family, such as GroES or an analogue thereof. A highly preferred analogue
 11 is the T4 polypeptide Gp31. GroES analogues, including Gp31, possess a mobile loop
 12 (Hunt, J. F., et al., (1997) *Cell* **90**, 361-371; Landry, S. J., et al., (1996) *Proc. Natl.
 13 Acad. Sci. U.S.A.* **93**, 11622-11627) which may be inserted into, or replaced, in order to
 14 fuse the polypeptide to the scaffold.

Description of Preferred Embodiments

Such a variant may be encoded by a related gene of the same gene family, by an allelic variant of a particular gene, or represent an alternative splicing variant of a molecular chaperone or foldase. Possible derivatives of the polypeptides employed in the invention

Actinomycetemcomitans; *Actinobacillus pleuropneumoniae*; *Aeromonas* *salmoneicida*; *Agrobacterium tumefaciens*; *Allochromatium vinosum*; *Amoeba proteus* *symbiotica* *bacterium*; *Aquifex aeolicus*; *Arabidopsis thaliana*; *Bacillus* *sp*; *Bacillus stearothermophilus*; *Bacillus subtilis*; *Baronella henselae*; *Bordetella pertussis*; *Borrelia burgdorferi*; *Brucella abortus*; *Buchnera aphidicola*; *Burkholderia cepacia* *Burkholderia vietnamensis*; *Campylobacter jejuni*; *Caulobacter crescentus*; *Chlamydia muridarium*; *Chlamydia trachomatis*; *Chlamydopila pneumoniae*; *Clostridium acetobutylicum*; *Clostridium perfringens*; *Clostridium thermocellum*; *Clostridium* *chaffeensis*; *Ehrlichia equi*; *Ehrlichia phagocytophila*; *Ehrlichia risticii*; *Ehrlichia sennetsu*; *Ehrlichia sp HGE agent*; *Enterobacter aerogenes*; *Enterobacter agglomerans* *Enterobacter amnigenus*; *Enterobacter asburiae*; *Enterobacter ergoviae*; *Enterobacter intermedius*; *Erwinia aphidicola*; *Erwinia carotovora*; *Erwinia herbicola*; *Escherichia* 30 *35*

20

Species:

Cp10 homologues are widespread throughout animals, plants and bacteria. For example, a search of GenBank indicates that cp10 homologues are known in the following

三

coli; *Francisella tularensis*; *Glycine max*; *Haemophilus ducreyi*; *Haemophilus influenzae Rd*; *Helicobacter pylori*; *Holospora obtusa*; *Homo sapiens*; *Klebsiella ornithinolytica*; *Klebsiella oxytoca*; *Klebsiella planticola*; *Klebsiella pneumoniae*; *Lactobacillus helveticus*; *Lactobacillus zae*; *Lactococcus lactis*; *Lawsonia intracellularis*; *Leptospira interrogans*; *Methylovorus sp strain SS*; *Mycobacterium avium*; *Mycobacterium avium subsp avium*; *Mycobacterium avium subsp paratuberculosis*; *Mycobacterium leprae*; *Mycobacterium tuberculosis*; *Mycoplasma genitalium*; *Mycoplasma pneumoniae*; *Myzus persicae primary endosymbiont*; *Neisseria gonorrhoeae*; *Oscillatoria sp NKBG*; *Pantoea ananas*; *Pasteurella multocida*; *Porphyromonas gingivalis*; *Pseudomonas aeruginosa*; *Pseudomonas aeruginosa*; *Pseudomonas putida*; *Rattus norvegicus*; *Rattus norvegicus*; *Rhizobium leguminosarum*; *Rhodobacter capsulatus*; *Rhodobacter sphaeroides*; *Rhodothermus marinus*; *Rickettsia prowazekii*; *Rickettsia rickettsii*; *Saccharomyces cerevisiae*; *Serratia ficaria*; *Serratia marcescens*; *Serratia rubidaea*; *Sinorhizobium meliloti*; *Sitophilus oryzae principal endosymbiont*; *Stenotrophomonas maltophilia*; *Streptococcus pneumoniae*; *Streptomyces albus*; *Streptomyces coelicolor*; *Streptomyces coelicolor*; *Streptomyces lividans*; *Synechococcus sp*; *Synechococcus vulgaris*; *Synechocystis sp*; *Thermoanaerobacter brockii*; *Thermotoga maritima*; *Thermus aquaticus*; *Treponema pallidum*; *Wolbachia sp*; *Zymomonas mobilis*.

20 An advantage of cpn10 family subunits is that they possess a mobile loop, responsible for the protein folding activity of the natural chaperonin, which may be removed without affecting the scaffold.

25 Cpn10 with a deleted mobile loop possesses no biological activity, making it an advantageously inert scaffold, thus minimising any potentially deleterious effects. Insertion of an appropriate biologically active polypeptide can confer a biological activity on the novel polypeptide thus generated. Indeed, the biological activity of the inserted polypeptide may be improved by incorporation of the biologically active polypeptide into the scaffold.

30 Alternative sites for peptide insertion are possible. An advantageous option is in the position equivalent to the roof beta hairpin in GroES. This involves replacement of Glu-60 in Gp31 by the desired peptide. The amino acid sequence is Pro(59)-Glu(60)-Gly(61).

produce libraries of polypeptide repositories, suitable for display and selection. As set forth therein, insertion of polypeptides may be combined with randomisation to 30 technique, including those set forth by Doi and Yamagawa (FEBS Letters (1999) 457:1-4). Figures 8 - 12 show various topologies and applications for scaffolded polypeptides in accordance with the present invention. In figures 8 and 9, the possible insertion sites for polypeptides are shown. Insertion of polypeptides may be performed by any suitable technique, including those set forth by Doi and Yamagawa (FEBS Letters (1999) 457:1-4).

25 Configurations of Oligomers according to the Invention

of these mutants for ring structures and would be suitable for use as scaffolds. 20 Mutation at Glu191-Gly, which restores activity by reducing the affinity for GroES. Both inactive *in vitro* because it is release to bind GroES. GroELSR2 has an additional between the two rings of GroEL (R452E, E461A, S463A and V464A) and is functionally GroEL (GroELSR1) contains four point mutations which effect the major attachment 25 mutations of cpn60 molecules may also be used. For example, the single ring mutant of

15 as that derived from *Mycobacterium tuberculosis*, is highly advantageous and confers an intended for vaccination purposes (see below), the use of an immunogenic scaffold, such may depend upon the intended application of the oligomer: for example, if the oligomer is 20 domains in order to impart mobility to the inserted polypeptide. The choice of scaffold *Cla* I sites) to maintain intact the hinge region between the equatorial and the apical particular between positions 197 and 333 (represented by *Sac*II engineered and unique 25 advantageously, polypeptides would be inserted between positions 191 and 376, in scaffolds. For example, the tetradecameric bacterial chaperonin GroEL may be used. Members of the cpn60/Hsp60 family of chaperonin molecules may also be used as 30 members. Inverse PCR may be used, to display the peptide on the opposite side of the scaffold. 35

5 Similarly, an insertion may be made at between positions 56 and 57 of the GroES Pro-Gly, leaving a blunt-ended restriction site for peptide insertion as a DNA fragment. This is conveniently converted to a Smal site at the DNA level (CCCGGG) encoding sequence, and at equivalent positions in other cpn60 family members. Alternatively, 10 sequence, and at equivalent positions in other cpn60 family members. Alternatively, Pro-Gly, leaving a blunt-ended restriction site for peptide insertion as a DNA fragment. This is conveniently converted to a Smal site at the DNA level (CCCGGG) encoding

Figure 10 illustrates the potential attachment sites for polypeptides to a circular scaffold, in this case Gp31. Reading from left to right, the figure shows: no attachment, attachment to the mobile loop, attachment to the roof β hairpin, attachment at both the mobile loop and the roof β hairpin, attachment at the C terminus, attachment at both N and C termini, 5 attachment at both N and C termini and the mobile loop, and attachment at both N and C termini, the roof β hairpin and the mobile loop. As will be apparent, further configurations are possible, and can be combined in any way in the heptamer, leading to a total of 5.4×10^8 possible configurations.

10 Recombinant DNA techniques

The present invention advantageously makes use of recombinant DNA technology in order to construct polypeptide monomers and oligomers. Advantageously, polypeptide monomers or oligomers may be expressed from nucleic acid sequences which encode 15 them.

As used herein, vector (or plasmid) refers to discrete elements that are used to introduce heterologous DNA into cells for either expression or replication thereof. Selection and use of such vehicles are well within the skill of the artisan. Many vectors are available, 20 and selection of appropriate vector will depend on the intended use of the vector, i.e. whether it is to be used for DNA amplification or for DNA expression, the size of the DNA to be inserted into the vector, and the host cell to be transformed with the vector. Each vector contains various components depending on its function (amplification of 25 DNA or expression of DNA) and the host cell for which it is compatible. The vector components generally include, but are not limited to, one or more of the following: an origin of replication, one or more marker genes, an enhancer element, a promoter, a transcription termination sequence and a signal sequence.

Both expression and cloning vectors generally contain nucleic acid sequence that enable 30 the vector to replicate in one or more selected host cells. Typically in cloning vectors, this sequence is one that enables the vector to replicate independently of the host chromosomal DNA, and includes origins of replication or autonomously replicating sequences. Such sequences are well known for a variety of bacteria, yeast and viruses.

The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria, the 2m plasmid origin is suitable for yeast, and various viral origins (e.g. SV 40, polyoma, adenovirus) are useful for cloning vectors in mammalian cells. Generally, the one class of organisms but can be transferred into another class of organisms for expression. For example, a vector is cloned in *E. coli* and then the same vector is transferred into yeast or mammalian cells even though it is not capable of replicating directly transferred into the host cells without any replication component.

Most expression vectors are shuttle vectors, i.e. they are capable of replication in at least one class of organisms but can be transferred into another class of organisms for expression. For example, a vector is cloned in *E. coli* and then the same vector is transferred into yeast or mammalian cells. This gene encodes a protein necessary for the survival of transformed host cells grown in a selective culture medium. Host cells not preferred to a selectable marker. This gene encodes a protein necessary for the survival of yeast of transformed host cells grown in a selective culture medium. Host cells not transformed with the vector containing the selection gene will not survive in the culture medium. Typical selection genes encode proteins that confer resistance to antibiotics and other toxins, e.g. ampicillin, neomycin, methotrexate or tetracycline, complement auxotrophic deficiencies, or supply critical nutrients not available from complex media.

As to a selective marker appropriate for yeast, any marker gene can be used which facilitates the selection for transformations due to the phenotypic expression of the marker gene. Suitable markers for yeast are, for example, those conferring resistance to antibiotics G418, hygromycin or bleomycin, or provide for prototrophy in an auxotrophic yeast mutant, for example the *URA3*, *LEU2*, *LYS2*, *TRP1*, or *HIS3* gene.

Since the replication of vectors is conveniently done in *E. coli*, an *E. coli* genetic marker or *pUC19*, which contain both *E. coli* replication origin and *E. coli* gene marker from *E. coli* plasmids, such as pBR322, Bluescript[®] vector or a *pUC* plasmid, e.g. *pUC18* and an *E. coli* origin of replication are advantageously included. These can be obtained 30 and an *E. coli* origin of replication are advantageously included. These can be obtained confirming resistance to antibiotics, such as ampicillin.

COS cells.

These are used in mammalian cells competent for high level DNA replication, such as origin of replication component is not needed for mammalian expression vectors unless polyoma, adenovirus) are useful for cloning vectors in mammalian cells. Generally, the 10 one class of organisms but can be transferred into another class of organisms for expression. For example, a vector is cloned in *E. coli* and then the same vector is transferred into yeast or mammalian cells even though it is not capable of replicating directly transferred into the host cells without any replication component.

Suitable selectable markers for mammalian cells are those that enable the identification of cells which have been transformed, such as dihydrofolate reductase (DHFR, methotrexate resistance), thymidine kinase, or genes conferring resistance to G418 or hygromycin. The 5 mammalian cell transformants are placed under selection pressure which only those transformants which have taken up and are expressing the marker are uniquely adapted to survive. In the case of a DHFR or glutamine synthase (GS) marker, selection pressure can be imposed by culturing the transformants under conditions in which the pressure is progressively increased, thereby leading to amplification (at its chromosomal integration 10 site) of both the selection gene and the linked DNA that encodes the polypeptide according to the invention. Amplification is the process by which genes in greater demand for the production of a protein critical for growth, together with closely associated genes which may encode a desired protein, are reiterated in tandem within the chromosomes of recombinant cells. Increased quantities of desired protein are usually 15 synthesised from thus amplified DNA.

Expression and cloning vectors usually contain a promoter that is recognised by the host organism and is operably linked to the heterologous nucleic acid coding sequence. Such a promoter may be inducible or constitutive. The promoters are operably linked to the 20 coding sequence by inserting the isolated promoter sequence into the vector. Many heterologous promoters may be used to direct amplification and/or expression of the coding sequence. The term "operably linked" refers to a juxtaposition wherein the components described are in a relationship permitting them to function in their intended manner. A control sequence "operably linked" to a coding sequence is ligated in such a 25 way that expression of the coding sequence is achieved under conditions compatible with the control sequences.

Promoters suitable for use with prokaryotic hosts include, for example, the β -lactamase and lactose promoter systems, alkaline phosphatase, the tryptophan (trp) promoter system 30 and hybrid promoters such as the tac promoter. Their nucleotide sequences have been published, thereby enabling the skilled worker operably to ligate them to the coding sequence, using linkers or adapters to supply any required restriction sites. Promoters for

5 Preferred expression vectors are bacterial expression vectors which comprise a promoter
of a bacteriophage such as phage or λ which is capable of functioning in the bacteria.
In one of the most widely used expression systems, the nucleic acid encoding the fusion
protein may be transcribed from the vector by λ RNA polymerase (Studier et al,
Methods in Enzymol, 185, 60-89, 1990). In the *E. coli* BL21(DE3) host strain, used in
conjunction with PET vectors, the λ RNA polymerase is produced from the λ -lysozyme
DE3 in the host bacterium, and its expression is under the control of the IPTG inducible
lac UV5 promoter. This system has been employed successfully for over-production of
many proteins. Alternatively the polymerase gene may be introduced on a lambda phage
by infection with an λ -phage such as the λ C6 phage which is commercially available
(Novagen, Madison, USA). other vectors include vectors containing the lambda PL
promoter such as PLEX (Invitrogen, NL), vectors containing the tac promoter such as
PTCHisXpressTM (Invitrogen) or PTc99 (Pharmacia Biotech, SE), or vectors containing
peptide may be recovered from the bacterial periplasmic space, or the culture medium, as
it will be produced as a soluble native peptide rather than in an inclusion body. The
sequence in order to facilitate secretion of the polypeptide from bacterial hosts, such that
suitable promoting sequences for use with yeast hosts may be regulated or constitutive
30 ADHII gene, the acid phosphatase (PH05) gene, a promoter of the TRLI gene, the ADHII or
Saccharomyces cerevisiae gene. Thus, the promoter of the TRLI gene, especially a
and are preferably derived from a highly expressed yeast gene, especially a
pheromone genes coding for the α - or α -factor or a promoter derived from a gene
encoding a glycolytic enzyme such as the enolase, glyceraldehyde-3-
phosphate dehydrogenase (GAP), 3-phosphoglycerate kinase (PGK), hexokinase,
pyruvate decarboxylase, phosphotuctokinase, glucose-6-phosphate isomerase, 3-

25

20 Moreover, the coding sequence according to the invention preferably includes a secretion
sequence in order to facilitate secretion of the polypeptide from bacterial hosts, such that
suitable promoting sequences for use with yeast hosts may be regulated or constitutive
appropiate.

Biolabs, MA, USA).
the tac promoter such as PK223-3 (Pharmacia Biotech) or PMAL (new England
PTCHisXpressTM (Invitrogen) or PTc99 (Pharmacia Biotech, SE), or vectors containing
promoter such as PLEX (Invitrogen, NL), vectors containing the tac promoter such as
(Novagen, Madison, USA). other vectors include vectors containing the lambda PL
promoter such as PLEX (Invitrogen, NL), vectors containing the tac promoter such as
PTCHisXpressTM (Invitrogen) or PTc99 (Pharmacia Biotech, SE), or vectors containing
peptide may be recovered from the bacterial periplasmic space, or the culture medium, as
it will be produced as a soluble native peptide rather than in an inclusion body. The
sequence in order to facilitate secretion of the polypeptide from bacterial hosts, such that
suitable promoting sequences for use with yeast hosts may be regulated or constitutive
30 ADHII gene, the acid phosphatase (PH05) gene, a promoter of the TRLI gene, the ADHII or
Saccharomyces cerevisiae gene. Thus, the promoter of the TRLI gene, the ADHII or
and are preferably derived from a highly expressed yeast gene, especially a
pheromone genes coding for the α - or α -factor or a promoter derived from a gene
encoding a glycolytic enzyme such as the enolase, glyceraldehyde-3-
phosphate dehydrogenase (GAP), 3-phosphoglycerate kinase (PGK), hexokinase,
pyruvate decarboxylase, phosphotuctokinase, glucose-6-phosphate isomerase, 3-

10

5 use in bacterial systems will also generally contain a Shine-Dalgarno sequence operably
linked to the coding sequence.

phosphoglycerate mutase, pyruvate kinase, triose phosphate isomerase, phosphoglucose isomerase or glucokinase genes, the *S. cerevisiae* GAL 4 gene, the *S. pombe* nmt 1 gene or a promoter from the TATA binding protein (TBP) gene can be used. Furthermore, it is possible to use hybrid promoters comprising upstream activation sequences (UAS) of one yeast gene and downstream promoter elements including a functional TATA box of another yeast gene, for example a hybrid promoter including the UAS(s) of the yeast PH05 gene and downstream promoter elements including a functional TATA box of the yeast GAP gene (PH05-GAP hybrid promoter). A suitable constitutive PH05 promoter is e.g. a shortened acid phosphatase PH05 promoter devoid of the upstream regulatory elements (UAS) such as the PH05 (-173) promoter element starting at nucleotide -173 and ending at nucleotide -9 of the PH05 gene.

Transcription from vectors in mammalian hosts may be controlled by promoters derived from the genomes of viruses such as polyoma virus, adenovirus, fowlpox virus, bovine papilloma virus, avian sarcoma virus, cytomegalovirus (CMV), a retrovirus and Simian Virus 40 (SV40), from heterologous mammalian promoters such as the actin promoter or a very strong promoter, e.g. a ribosomal protein promoter, provided such promoters are compatible with the host cell systems.

Transcription of a coding sequence by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are relatively orientation and position independent. Many enhancer sequences are known from mammalian genes (e.g. elastase and globin). However, typically one will employ an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270) and the CMV early promoter enhancer. The enhancer may be spliced into the vector at a position 5' or 3' to the coding sequence, but is preferably located at a site 5' from the promoter.

Advantageously, a eukaryotic expression vector may comprise a locus control region (LCR). LCRs are capable of directing high-level integration site independent expression of transgenes integrated into host cell chromatin, which is of importance especially where the coding sequence is to be expressed in the context of a permanently-transfected

eukaryotic cell line in which chromosomal integration of the vector has occurred, in vectors designed for gene therapy applications or in transgenic animals.

An expression vector includes any vector capable of expressing nucleic acids that are operatively linked with regulatory sequences, such as promoter regions, that are capable of expression of such DNAs. Thus, an expression vector refers to a recombinant DNA or RNA construct, such as a plasmid, a phage, recombinant virus or other vector, that upon introduction into an appropriate host cell, results in expression of the cloned DNA. Appropriate expression vectors are well known to those with ordinary skill in the art and include those that are replicable in eukaryotic and/or prokaryotic cells and those that remain episomal or those which integrate into the host cell genome. For example, nucleic acids may be inserted into a vector suitable for expression of DNAs in mammalian cells, e.g. a CMV enhancer-based vector such as PEVRF (Mathias, et al., (1989) NAR 17, 6418).

An expression vector includes any vector capable of expressing nucleic acids that are operationally linked with regulatory sequences, such as promoter regions, that are capable of expression of such DNAs. Thus, an expression vector refers to a recombinant DNA or RNA construct, such as a plasmid, a phage, recombinant virus or other vector, that upon introduction into an appropriate host cell, results in expression of the cloned DNA. Appropriate expression vectors are well known to those with ordinary skill in the art and include those that are replicable in eukaryotic and/or prokaryotic cells and those that remain episomal or those which integrate into the host genome. For example, nucleic acids may be inserted into a vector suitable for expression of DNAs in mammalian cells, e.g. a CMV enhancer-based vector such as PEVRF (Mathias, et al., (1989) NAR 17,

Construction of vectors according to the invention employs conventional ligation techniques. Isolated plasmids or DNA fragments are cleaved, tailored, and religated in the form desired to generate the plasmids required. If desired, analyses to confirm correct sequences in the constructed plasmids is performed in a known fashion. Suitable methods for constructing expression vectors, preparing *in vitro* transcripts, introducing DNA into host cells, and performing analyses for assessing expression and function are known to those skilled in the art. Gene presence, amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA, dot blotting (DNA or RNA analysis), or *in situ* hybridization, using an appropriately labelled probe which may be based on a sequence provided herein. Those skilled in the art will readily envisage how these methods may be used.

Description of Preferred Embodiments

In a preferred aspect, the present invention relates to a method for oligomerising 5 polypeptides and novel oligomeric polypeptides producible thereby. By incorporating polypeptides into a scaffold which is capable of oligomerising, as described herein, oligomers may be produced wherein the chosen polypeptides are juxtaposed. The polypeptides selected for oligomerisation may be the same or different; thus, it is possible to produce homooligomers or heterooligomers.

10

Preferably, the oligomeric proteins of the invention are oligomeric minichaperone polypeptides. Oligomeric minichaperones have been shown to possess particularly advantageous properties.

15

The invention thus relates to the use of oligomeric molecular chaperone fragments, whether alone or in combination with other polypeptides, for potentiating the folding or refolding of polypeptides.

20

The present invention may be practised in a number of configurations, according to the required use to which the invention is to be put. In a first configuration, the invention relates to the use of minichaperone oligomers alone to potentiate the folding or refolding of polypeptides. This may be performed *in vivo* or *in vitro*, in solution or on a solid support. For example, oligomerised minichaperones may be immobilised onto resins and packed into columns for use in refolding polypeptides which are passed through the 25 column. Methods for immobilising minichaperones are described in International patent application WO99/05163, incorporated herein by reference.

30

In an alternative configuration, minichaperone oligomers according to the invention may be expressed *in vivo* or administered to cells or organisms *in vivo* in order to potentiate protein folding therein.

In a second configuration, the invention provides a combination of a molecular chaperone and a thiol/disulphide oxidoreductase to facilitate protein folding. One or both of the

chaperones and thiol/disulphide oxidoreductases may be incorporated into the scaffold according to the invention. The combination of a molecular chaperone and a thiol/disulphide oxidoreductase provides a synergistic effect on protein folding which results in a greater quantity of active, correctly folded protein being produced than would be expected from a merely additive relationship. Advantageously, one or more of the components used to potentiate protein folding in accordance with the present invention is immobilised on a solid support. However, both molecular chaperones and thiol/disulphide oxidoreductases may be used in solution. They may be used in free solution, but also in suspension, for example bound to a matrix such as beads, for example Sepharose beads, or bound to solid surfaces which are in contact with solutions, such as the inside surfaces of bottles containing solutions, test tubes and the like.

In a third configuration, the invention relates to the use of a combination of a molecular chaperone and a thiol/disulphide oxidoreductase with a peptidyl prolyl isomerase. The peptidyl prolyl isomerase may be present either incorporated into the oligomeric scaffold, bound to a solid support, or in solution. Moreover, it may be bound to beads suspended in solution. The peptidyl prolyl isomerases may be used together with a molecular chaperone alone, with a thiol/disulphide oxidoreductase alone, or with both a molecular chaperone and a thiol/disulphide oxidoreductase. In the latter case, further synergistic effects are apparent over the additive effects which would be expected from the use of the three components together. In particular, an increase in the proportion of the folded protein which is recovered as monodisperse protein, as opposed to aggregated protein, increases substantially.

Used in accordance with any of the foregoing configurations, or otherwise in accordance with the following claims, the invention may be used to facilitate protein folding in a variety of situations. For example, the invention may be used to assist in refolding recombinantly produced polypeptides, which are obtained in an unfolded or misfolded form. Thus, recombinantly produced polypeptides may be passed down a column on which is immobilised a composition comprising protein disulphide isomerase and/or a molecular chaperone and/or a prolyl peptidyl isomerase.

30

20

15

10

5

In an alternative embodiment, in a the invention may be employed to maintain the folded conformation of proteins, for example during storage, in order to increase shelf life. under storage conditions, many proteins lose their activity, as a result of disruption of correct folding. The presence of molecular chaperones, in combination with foldases, 5 reduces or reverses the tendency of polypeptides to become unfolded and thus greatly increases the shelf life thereof. In this embodiment, the invention may be applied to reagents which comprise polypeptide components, such as enzymes, tissue culture components, and other proteinaceous reagents stored in solution.

10 In a further embodiment, the invention may be used to potentiate the correct folding of proteins which, through storage, exposure to denaturing conditions or otherwise, have become misfolded. Thus, the invention may be used to recondition reagents or other proteins. For example, proteins in need of reconditioning may be passed down a column to which is immobilised a combination of reagents in accordance with the invention.

15 Alternatively, beads having immobilised thereon such a combination may be suspended in a solution comprising the proteins in need of reconditioning. Moreover, the components of the combination according to the invention may be added in solution to the proteins in need of reconditioning.

20 As noted above, the components of the combination according to the invention may comprise derivatives of molecular chaperones or foldases, including variants of such polypeptides which retain common structural features thereof. Variants which retain common structural features can be fragments of molecular chaperones or foldases. Fragments of molecular chaperones or foldases comprise smaller polypeptides derived 25 from therefrom. Preferably, smaller polypeptides derived from the molecular chaperones or foldases according to the invention define a single feature which is characteristic of the molecular chaperones or foldases. Fragments may in theory be almost any size, as long as they retain the activity of the molecular chaperones or foldases described herein.

30 With respect to molecular chaperones of the GroEL/hsp-60 family, a preferred set of fragments have been identified which possess the desired activity. These fragments are set forth in our copending international patent application PCT/GB96/02980 and in essence comprise any fragment comprising at least amino acid residues 230-271 of intact

crossreactivity with the native forms of the relevant molecular chaperone or foldase. foldases can be prepared by recombinant methods and screened for immunoreactivity, for example, substitutional, deletional or insertion of molecular chaperones or resulting e.g. in an addition, exchange and/or deletion of one or more amino acids. For molecules comprising by the invention. Mutants may be produced from a DNA encoding a foldases may moreover be made to the fragments of the molecular chaperones or substitutions may be made to the molecular chaperones or chaperones or foldases, as may truncations from the 5' or 3' ends. Deletions and substitutions may be made substantially without altering the nature of the molecular chaperones or foldases described herein. Thus, conservative amino acid deletions, additions or substitutions, subject to the requirement to maintain the activity of including mutants of fragments and other derivatives, which may contain amino acid derivatives of the molecular chaperones or foldases also comprise mutants thereof, 20 25 30

residues 191-376, 191-345 and 191-335 of the sequence of intact GroEL. In a highly preferred embodiment, the fragments are selected from the group consisting of

15 apical domain of GroEL as defined above. Preferably, molecular chaperones according to the invention are homologous to, or are capable of hybridizing under stringent conditions with, a region corresponding to the apical domain spans amino acids 191-376 of intact GroEL. This domain is found to be other molecular chaperones, or a region homologous thereto as defined herein. The apical domain sequence as set forth above; in other words, they must not have an Alanine hsp-60 chaperone. Where the fragments are GroEL fragments, they must not possess the 10 advantageously, the fragments comprise the apical domain of GroEL, or its equivalent in other molecular chaperones, or a region homologous thereto as defined herein. The apical domain sequence as set forth above; in other words, they must not have an Alanine

homologous amongst a wide number of species and chaperone types. 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195 200 205 210 215 220 225 230 235 240 245 250 255 260 265 270 275 280 285 290 295 300 305 310 315 320 325 330 335 340 345 350 355 360 365 370 375 380 385 390 395 400 405 410 415 420 425 430 435 440 445 450 455 460 465 470 475 480 485 490 495 500 505 510 515 520 525 530 535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 675 680 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800 805 810 815 820 825 830 835 840 845 850 855 860 865 870 875 880 885 890 895 900 905 910 915 920 925 930 935 940 945 950 955 960 965 970 975 980 985 990 995 1000 1005 1010 1015 1020 1025 1030 1035 1040 1045 1050 1055 1060 1065 1070 1075 1080 1085 1090 1095 1100 1105 1110 1115 1120 1125 1130 1135 1140 1145 1150 1155 1160 1165 1170 1175 1180 1185 1190 1195 1200 1205 1210 1215 1220 1225 1230 1235 1240 1245 1250 1255 1260 1265 1270 1275 1280 1285 1290 1295 1300 1305 1310 1315 1320 1325 1330 1335 1340 1345 1350 1355 1360 1365 1370 1375 1380 1385 1390 1395 1400 1405 1410 1415 1420 1425 1430 1435 1440 1445 1450 1455 1460 1465 1470 1475 1480 1485 1490 1495 1500 1505 1510 1515 1520 1525 1530 1535 1540 1545 1550 1555 1560 1565 1570 1575 1580 1585 1590 1595 1600 1605 1610 1615 1620 1625 1630 1635 1640 1645 1650 1655 1660 1665 1670 1675 1680 1685 1690 1695 1700 1705 1710 1715 1720 1725 1730 1735 1740 1745 1750 1755 1760 1765 1770 1775 1780 1785 1790 1795 1800 1805 1810 1815 1820 1825 1830 1835 1840 1845 1850 1855 1860 1865 1870 1875 1880 1885 1890 1895 1900 1905 1910 1915 1920 1925 1930 1935 1940 1945 1950 1955 1960 1965 1970 1975 1980 1985 1990 1995 2000 2005 2010 2015 2020 2025 2030 2035 2040 2045 2050 2055 2060 2065 2070 2075 2080 2085 2090 2095 2100 2105 2110 2115 2120 2125 2130 2135 2140 2145 2150 2155 2160 2165 2170 2175 2180 2185 2190 2195 2200 2205 2210 2215 2220 2225 2230 2235 2240 2245 2250 2255 2260 2265 2270 2275 2280 2285 2290 2295 2300 2305 2310 2315 2320 2325 2330 2335 2340 2345 2350 2355 2360 2365 2370 2375 2380 2385 2390 2395 2400 2405 2410 2415 2420 2425 2430 2435 2440 2445 2450 2455 2460 2465 2470 2475 2480 2485 2490 2495 2500 2505 2510 2515 2520 2525 2530 2535 2540 2545 2550 2555 2560 2565 2570 2575 2580 2585 2590 2595 2600 2605 2610 2615 2620 2625 2630 2635 2640 2645 2650 2655 2660 2665 2670 2675 2680 2685 2690 2695 2700 2705 2710 2715 2720 2725 2730 2735 2740 2745 2750 2755 2760 2765 2770 2775 2780 2785 2790 2795 2800 2805 2810 2815 2820 2825 2830 2835 2840 2845 2850 2855 2860 2865 2870 2875 2880 2885 2890 2895 2900 2905 2910 2915 2920 2925 2930 2935 2940 2945 2950 2955 2960 2965 2970 2975 2980 2985 2990 2995 3000 3005 3010 3015 3020 3025 3030 3035 3040 3045 3050 3055 3060 3065 3070 3075 3080 3085 3090 3095 3100 3105 3110 3115 3120 3125 3130 3135 3140 3145 3150 3155 3160 3165 3170 3175 3180 3185 3190 3195 3200 3205 3210 3215 3220 3225 3230 3235 3240 3245 3250 3255 3260 3265 3270 3275 3280 3285 3290 3295 3300 3305 3310 3315 3320 3325 3330 3335 3340 3345 3350 3355 3360 3365 3370 3375 3380 3385 3390 3395 3400 3405 3410 3415 3420 3425 3430 3435 3440 3445 3450 3455 3460 3465 3470 3475 3480 3485 3490 3495 3500 3505 3510 3515 3520 3525 3530 3535 3540 3545 3550 3555 3560 3565 3570 3575 3580 3585 3590 3595 3600 3605 3610 3615 3620 3625 3630 3635 3640 3645 3650 3655 3660 3665 3670 3675 3680 3685 3690 3695 3700 3705 3710 3715 3720 3725 3730 3735 3740 3745 3750 3755 3760 3765 3770 3775 3780 3785 3790 3795 3800 3805 3810 3815 3820 3825 3830 3835 3840 3845 3850 3855 3860 3865 3870 3875 3880 3885 3890 3895 3900 3905 3910 3915 3920 3925 3930 3935 3940 3945 3950 3955 3960 3965 3970 3975 3980 3985 3990 3995 4000 4005 4010 4015 4020 4025 4030 4035 4040 4045 4050 4055 4060 4065 4070 4075 4080 4085 4090 4095 4100 4105 4110 4115 4120 4125 4130 4135 4140 4145 4150 4155 4160 4165 4170 4175 4180 4185 4190 4195 4200 4205 4210 4215 4220 4225 4230 4235 4240 4245 4250 4255 4260 4265 4270 4275 4280 4285 4290 4295 4300 4305 4310 4315 4320 4325 4330 4335 4340 4345 4350 4355 4360 4365 4370 4375 4380 4385 4390 4395 4400 4405 4410 4415 4420 4425 4430 4435 4440 4445 4450 4455 4460 4465 4470 4475 4480 4485 4490 4495 4500 4505 4510 4515 4520 4525 4530 4535 4540 4545 4550 4555 4560 4565 4570 4575 4580 4585 4590 4595 4600 4605 4610 4615 4620 4625 4630 4635 4640 4645 4650 4655 4660 4665 4670 4675 4680 4685 4690 4695 4700 4705 4710 4715 4720 4725 4730 4735 4740 4745 4750 4755 4760 4765 4770 4775 4780 4785 4790 4795 4800 4805 4810 4815 4820 4825 4830 4835 4840 4845 4850 4855 4860 4865 4870 4875 4880 4885 4890 4895 4900 4905 4910 4915 4920 4925 4930 4935 4940 4945 4950 4955 4960 4965 4970 4975 4980 4985 4990 4995 5000 5005 5010 5015 5020 5025 5030 5035 5040 5045 5050 5055 5060 5065 5070 5075 5080 5085 5090 5095 5100 5105 5110 5115 5120 5125 5130 5135 5140 5145 5150 5155 5160 5165 5170 5175 5180 5185 5190 5195 5200 5205 5210 5215 5220 5225 5230 5235 5240 5245 5250 5255 5260 5265 5270 5275 5280 5285 5290 5295 5300 5305 5310 5315 5320 5325 5330 5335 5340 5345 5350 5355 5360 5365 5370 5375 5380 5385 5390 5395 5400 5405 5410 5415 5420 5425 5430 5435 5440 5445 5450 5455 5460 5465 5470 5475 5480 5485 5490 5495 5500 5505 5510 5515 5520 5525 5530 5535 5540 5545 5550 5555 5560 5565 5570 5575 5580 5585 5590 5595 5600 5605 5610 5615 5620 5625 5630 5635 5640 5645 5650 5655 5660 5665 5670 5675 5680 5685 5690 5695 5700 5705 5710 5715 5720 5725 5730 5735 5740 5745 5750 5755 5760 5765 5770 5775 5780 5785 5790 5795 5800 5805 5810 5815 5820 5825 5830 5835 5840 5845 5850 5855 5860 5865 5870 5875 5880 5885 5890 5895 5900 5905 5910 5915 5920 5925 5930 5935 5940 5945 5950 5955 5960 5965 5970 5975 5980 5985 5990 5995 6000 6005 6010 6015 6020 6025 6030 6035 6040 6045 6050 6055 6060 6065 6070 6075 6080 6085 6090 6095 6100 6105 6110 6115 6120 6125 6130 6135 6140 6145 6150 6155 6160 6165 6170 6175 6180 6185 6190 6195 6200 6205 6210 6215 6220 6225 6230 6235 6240 6245 6250 6255 6260 6265 6270 6275 6280 6285 6290 6295 6300 6305 6310 6315 6320 6325 6330 6335 6340 6345 6350 6355 6360 6365 6370 6375 6380 6385 6390 6395 6400 6405 6410 6415 6420 6425 6430 6435 6440 6445 6450 6455 6460 6465 6470 6475 6480 6485 6490 6495 6500 6505 6510 6515 6520 6525 6530 6535 6540 6545 6550 6555 6560 6565 6570 6575 6580 6585 6590 6595 6600 6605 6610 6615 6620 6625 6630 6635 6640 6645 6650 6655 6660 6665 6670 6675 6680 6685 6690 6695 6700 6705 6710 6715 6720 6725 6730 6735 6740 6745 6750 6755 6760 6765 6770 6775 6780 6785 6790 6795 6800 6805 6810 6815 6820 6825 6830 6835 6840 6845 6850 6855 6860 6865 6870 6875 6880 6885 6890 6895 6900 6905 6910 6915 6920 6925 6930 6935 6940 6945 6950 6955 6960 6965 6970 6975 6980 6985 6990 6995 7000 7005 7010 7015 7020 7025 7030 7035 7040 7045 7050 7055 7060 7065 7070 7075 7080 7085 7090 7095 7100 7105 7110 7115 7120 7125 7130 7135 7140 7145 7150 7155 7160 7165 7170 7175 7180 7185 7190 7195 7200 7205 7210 7215 7220 7225 7230 7235 7240 7245 7250 7255 7260 7265 7270 7275 7280 7285 7290 7295 7300 7305 7310 7315 7320 7325 7330 7335 7340 7345 7350 7355 7360 7365 7370 7375 7380 7385 7390 7395 7400 7405 7410 7415 7420 7425 7430 7435 7440 7445 7450 7455 7460 7465 7470 7475 7480 7485 7490 7495 7500 7505 7510 7515 7520 7525 7530 7535 7540 7545 7550 7555 7560 7565 7570 7575 7580 7585 7590 7595 7600 7605 7610 7615 7620 7625 7630 7635 7640 7645 7650 7655 7660 7665 7670 7675 7680 7685 7690 7695 7700 7705 7710 7715 7720 7725 7730 7735 7740 7745 7750 7755 7760 7765 7770 7775 7780 7785 7790 7795 7800 7805 7810 7815 7820 7825 7830 7835 7840 7845 7850 7855 7860 7865 7870 7875 7880 7885 7890 7895 7900 7905 7910 7915 7920 7925 7930 7935 7940 7945 7950 7955 7960 7965 7970 7975 7980 7985 7990 7995 8000 8005 8010 8015 8020 8025 8030 8035 8040 8045 8050 8055 8060 8065 8070 8075 8080 8085 8090 8095 8100 8105 8110 8115 8120 8125 8130 8135 8140 8145 8150 8155 8160 8165 8170 8175 8180 8185 8190 8195 8200 8205 8210 8215 8220 8225 8230 8235 8240 8245 8250 8255 8260 8265 8270 8275 8280 8285 8290 8295 8300 8305 8310 8315 8320 8325 8330 8335 8340 8345 8350 8355 8360 8365 8370 8375 8380 8385 8390 8395 8400 8405 8410 8415 8420 8425 8430 8435 8440 8445 8450 8455 8460 8465 8470 8475 8480 8485 8490 8495 8500 8505 8510 8515 8520 8525 8530 8535 8540 8545 8550 8555 8560 8565 8570 8575 8580 8585 8590 8595 8600 8605 8610 8615 8620 8625 8630 8635 8640 8645 8650 8655 8660 8665 8670 8675 8680 8685 8690 8695 8700 8705 8710 8715 8720 8725 8730 8735 8740 8745 8750 8755 8760 8765 8770 8775 8780 8785 8790 8795 8800 8805 8810 8815 8820 8825 8830 8835 8840 8845 8850 8855 8860 8865 8870 8875 8880 8885 8890 8895 8900 8905 8910 8915 8920 8925 8930 8935 8940 8945 8950 8955 8960 8965 8970 8975 8980 8985 8990 8995 9000 9005 9010 9015 9020 9025 9030 9035 9040 9045 9050 9055 9060 9065 9070 9075 9080 9085 9090 9095 9100 9105 9110 9115 9120 9125 9130 9135 9140 9145 9150 9155 9160 9165 9170 9175 9180 9185 9190 9195 9200 9205 9210 9215 9220 9225 9230 9235 9240 9245 9250 9255 9260 9265 9270 9275 9280 9285 9290 9295 9300 9305 9310 9315 9320 9325 9330 9335 9340 9345 9350 9355 9360 9365 9370 9375 9380 9385 9390 9395 9400 9405 9410 9415 9420 9425 9430 9435 9440 9445 9450 9455 9460 9465 9470 9475 9480 9485 9490 9495 9500 9505 9510 9515 9520 9525 9530 9535 9540 9545 9550 9555 9560 9565 9570 9575 9580 9585 9590 9595 9600 9605 9610 9615 9620 9625 9630 9635 9640 9645 9650 9655 9660 9665 9670 9675 9680 9685 9690 9695 9700 9705 9710 9715 9720 9725 9730 9735 9740 9745 9750 9755 9760 9765 9770 9775 9780 9785 9790 9795 9800 9805 9810 9815 9820 9825 9830 9835 9840 9845 9850 9855 9860 9865 9870 9875 9880 9885 9890 9895 9900 9905 9910 9915 9920 9925 9930 9935 9940 9945 9950 9955 9960 9965 9970 9975 9980 9985 9990 9995 10000 10005 10010 10015 10020 10025 10030 10035 10040 10045 10050 10055 10060 10065 10070 10075 10080 10085 10090 10095 10100 10105 10110 10115 10120 10125 10130 10135 10140 10145 10150 10155 10160 10165 10170 10175 10180 10185 10190 10195 10200 10205 10210 10215 10220 10225 10230 10235 10240 10245 10250 10255 10260 10265 10270 10275 10280 10285 10290 10295 10300 10305 10310 10315 10320 10325 10330 10335 10340 10345 10350 10355 10360 10365 10370 10375 10380 10385 10390 10395 10400 10405 10410 10415 10420 10425 10430 10435 10440 10445 10450 10455 10460 10465 10470 10475 10480 10485 10490 10495 10500 10505 10510 10515 10520 10525 10530 10535 10540 10545 10550 10555 10560 10565 10570 10575 10580 10585 10590 10595 10600 10605 10610 10615 10620 10625 10630 10635 10640 10645 10650 10655 10660 10665 10670 10675 10680 10685 10690 10695 10700 10705 10710 10715 10720 10725 10730 10735 10740 10745 10750 10755 10760 10765 10770 10775 10780 10785 10790 10795 10800 10805 10810 10815 10820 10825 10830 10835 10840 10845

The fragments, mutants and other derivative of the molecular chaperones or foldases preferably retain substantial homology with the native molecular chaperones or foldases. As used herein, "homology" means that the two entities share sufficient characteristics for the skilled person to determine that they are similar in origin and function. Preferably, 5 homology is used to refer to sequence identity. Thus, the derivatives of molecular chaperones or foldases preferably retain substantial sequence identity with native forms of the relevant molecular chaperone or foldase.

In the context of the present invention, a homologous sequence is taken to include an 10 amino acid sequence which is at least 60, 70, 80 or 90% identical, preferably at least 95 or 98% identical at the amino acid level over at least 20, preferably 30 amino of the minichaperone. In particular, homology should typically be considered with respect to those regions of the sequence known to be essential for the chaperone activity rather than 15 non-essential neighbouring sequences. Although homology can also be considered in terms of similarity (i.e. amino acid residues having similar chemical properties/functions), in the context of the present invention it is preferred to express homology in terms of sequence identity.

Homology comparisons can be conducted by eye, or more usually, with the aid of readily 20 available sequence comparison programs. These commercially available computer programs can calculate % homology between two or more sequences.

% homology may be calculated over contiguous sequences, i.e. one sequence is aligned with the other sequence and each amino acid in one sequence directly compared with the 25 corresponding amino acid in the other sequence, one residue at a time. This is called an "ungapped" alignment. Typically, such ungapped alignments are performed only over a relatively short number of residues (for example less than 50 contiguous amino acids).

Although this is a very simple and consistent method, it fails to take into consideration that, 30 for example, in an otherwise identical pair of sequences, one insertion or deletion will cause the following amino acid residues to be put out of alignment, thus potentially resulting in a large reduction in % homology when a global alignment is performed. Consequently, most sequence comparison methods are designed to produce optimal alignments that take into

5 However, these more complex methods assign „gap penalties“ to each gap that occurs in the alignment so that, for the same number of identical amino acids, a sequence alignment with as few gaps as possible - reflecting higher relatedness between the two compared sequences - will achieve a higher score than one with many gaps. „Affine gap costs“ are typically used that charge a relatively high cost for the existence of a gap and a smaller penalty for each subsequent residue in the gap. This is the most commonly used gap scoring system. High gap penalties will of course produce optimised alignments with fewer gaps. Most alignments taking into consideration gap penalties. A suitable computer program for calculating maximum % homology therefore firstly requires the production of an optimal alignment, taking into consideration gap penalties. A suitable computer program for calculating maximum % homology therefore firstly requires the production of an optimal alignment.

10 values or a custom symbol comparison table if supplied (see user manual for further details of programs. GCG Wisconsin programs generally use either the public default matrix commonly used is the BLOSUM62 matrix - the default matrix for the BLAST process itself is typically not based on an all-or-nothing pair comparison. Instead, a comparison based on chemical similarity or evolutionary distance. An example of such a scaled similarity score matrix is generally used that assigns scores to each pairwise comparison based on chemical similarity or evolutionary distance. Instead, a process itself is typically not based on an all-or-nothing pair comparison. Instead, a

15 Although the final % homology can be measured in terms of identity, the alignment however it is preferred to use the GCG Bestfit program. GCG Wisconsin Bestfit package (see <http://www.ncbi.nlm.nih.gov/BLAST/>) is available for sequence searching at, for example, <http://www.2.ebi.ac.uk/blast3>) and the GENEWORKS suite of comparison tools. MOL. Biol., 403-410; FASTA is available for online searching at, for example, <http://www.ncbi.nlm.nih.gov/BLAST/>, FASTA (Atschul *et al.*, 1990, J. BLAST package (see <http://www.ncbi.nlm.nih.gov/BLAST/>). Examples of other software than can perform sequence comparisons include, but are not limited to, the Wisconsin, U.S.A.; Devereux *et al.*, 1984, Nucleic Acids Research 12:387). Examples of carrying out such an alignment is the GCG Wisconsin Bestfit package (University of Wisconsin, U.S.A.; Devereux *et al.*, 1984, Nucleic Acids Research 12:387). Examples of alignment, taking into consideration gap penalties. A suitable computer program for calculating maximum % homology therefore firstly requires the production of an optimal alignment.

20 25 Although the final % homology can be measured in terms of identity, the alignment however it is preferred to use the GCG Bestfit program. GCG Wisconsin Bestfit package (see <http://www.ncbi.nlm.nih.gov/BLAST/>) is available for sequence searching at, for example, <http://www.2.ebi.ac.uk/blast3>) and the GENEWORKS suite of comparison tools. MOL. Biol., 403-410; FASTA is available for online searching at, for example, <http://www.ncbi.nlm.nih.gov/BLAST/>, FASTA (Atschul *et al.*, 1990, J. BLAST package (see <http://www.ncbi.nlm.nih.gov/BLAST/>). Examples of other software than can perform sequence comparisons include, but are not limited to, the Wisconsin, U.S.A.; Devereux *et al.*, 1984, Nucleic Acids Research 12:387). Examples of carrying out such an alignment is the GCG Wisconsin Bestfit package (University of Wisconsin, U.S.A.; Devereux *et al.*, 1984, Nucleic Acids Research 12:387). Examples of alignment, taking into consideration gap penalties. A suitable computer program for calculating maximum % homology therefore firstly requires the production of an optimal alignment.

30 35 However, it is preferred to use the GCG Bestfit program. GCG Wisconsin Bestfit package (see below) the default gap penalty for amino acid values when using such software for sequence comparisons. For example when using the programs allow the gap penalties to be modified. However, it is preferred to use the default gap penalties will of course produce optimised alignments with fewer gaps. Most alignments carrying out such an alignment is the GCG Wisconsin Bestfit package (University of Wisconsin, U.S.A.; Devereux *et al.*, 1984, Nucleic Acids Research 12:387). Examples of alignment, taking into consideration gap penalties. A suitable computer program for calculating maximum % homology therefore firstly requires the production of an optimal alignment.

40 45 However, these more complex methods assign „gap penalties“ to each gap that occurs in the alignment so that, for the same number of identical amino acids, a sequence alignment with as few gaps as possible - reflecting higher relatedness between the two compared sequences - will achieve a higher score than one with many gaps. „Affine gap costs“ are typically used that charge a relatively high cost for the existence of a gap and a smaller penalty for each subsequent residue in the gap. This is the most commonly used gap scoring system. High gap penalties will of course produce optimised alignments with fewer gaps. Most alignments carrying out such an alignment is the GCG Wisconsin Bestfit package (University of Wisconsin, U.S.A.; Devereux *et al.*, 1984, Nucleic Acids Research 12:387). Examples of alignment, taking into consideration gap penalties. A suitable computer program for calculating maximum % homology therefore firstly requires the production of an optimal alignment.

50 55 However, these more complex methods assign „gap penalties“ to each gap that occurs in the alignment so that, for the same number of identical amino acids, a sequence alignment with as few gaps as possible - reflecting higher relatedness between the two compared sequences - will achieve a higher score than one with many gaps. „Affine gap costs“ are typically used that charge a relatively high cost for the existence of a gap and a smaller penalty for each subsequent residue in the gap. This is the most commonly used gap scoring system. High gap penalties will of course produce optimised alignments with fewer gaps. Most alignments carrying out such an alignment is the GCG Wisconsin Bestfit package (University of Wisconsin, U.S.A.; Devereux *et al.*, 1984, Nucleic Acids Research 12:387). Examples of alignment, taking into consideration gap penalties. A suitable computer program for calculating maximum % homology therefore firstly requires the production of an optimal alignment.

details). It is preferred to use the public default values for the GCG package, or in the case of other software, the default matrix, such as BLOSUM62.

Once the software has produced an optimal alignment, it is possible to calculate %
5 homology, preferably % sequence identity. The software typically does this as part of the sequence comparison and generates a numerical result.

Alternatively, sequence similarity may be defined according to the ability to hybridise to a complementary strand of a chaperone or foldase sequence as set forth above.

10

Preferably, the sequences are able to hybridise with high stringency. Stringency of hybridisation refers to conditions under which polynucleic acids hybrids are stable. Such conditions are evident to those of ordinary skill in the field. As known to those of skill in the art, the stability of hybrids is reflected in the melting temperature (Tm) of the hybrid
15 which decreases approximately 1 to 1.5°C with every 1% decrease in sequence homology. In general, the stability of a hybrid is a function of sodium ion concentration and temperature. Typically, the hybridisation reaction is performed under conditions of higher stringency, followed by washes of varying stringency.

20 As used herein, high stringency refers to conditions that permit hybridisation of only those nucleic acid sequences that form stable hybrids in 1 M Na+ at 65-68 °C. High stringency conditions can be provided, for example, by hybridisation in an aqueous solution containing 6x SSC, 5x Denhardt's, 1 % SDS (sodium dodecyl sulphate), 0.1 Na+ pyrophosphate and 0.1 mg/ml denatured salmon sperm DNA as non specific competitor.
25 Following hybridisation, high stringency washing may be done in several steps, with a final wash (about 30 min) at the hybridisation temperature in 0.2 - 0.1x SSC, 0.1 % SDS.

30 Moderate stringency refers to conditions equivalent to hybridisation in the above described solution but at about 60-62°C. In that case the final wash is performed at the hybridisation temperature in 1x SSC, 0.1 % SDS.

Low stringency refers to conditions equivalent to hybridization in the above described solution at about 50-52°C. In that case, the final wash is performed at the hybridization temperature in 2x SSC, 0.1% SDS.

It is understood that these conditions may be adapted and duplicated using a variety of buffers, e.g. formamide-based buffers, and temperatures. Denhardt's solution and SSC are well known to those of skill in the art as are other suitable hybridization buffers (see, e.g. Sambrook, et al., eds. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, New York or Ausubel, et al., eds. (1990) *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc.). Optimal hybridization conditions have to be determined empirically, as the length and the GC content of the probe also play a role.

The invention also envisages the administration of polypeptide oligomers according to the invention as compositions, preferably for the treatment of diseases associated with protein misfolding. The active compound may be administered in a convenient manner such as by the oral, intravenous (where water soluble), intramuscular, subcutaneous, intranasal, intradermal or suppository routes or implanting (e.g. using slow release molecules). Depending on the route of administration, the active ingredient may be required to be coated in a material to protect said ingredients from the action of enzymes, acids and other natural conditions which may inactivate said ingredient.

In order to administer the combination by other than parenteral administration, it will be coated by, or administered with, a material to prevent its inactivation. For example, the combination may be administered in an adjuvant, co-administered with enzyme inhibitors or in liposomes. Adjuvant is used in its broadest sense and includes any immune stimulating compound such as interferon. Adjuvants contemplated herein include resorcinols, non-ionic surfactants such as polyoxyethylene oleyl ether and n-hexadecyl polyethylene ether. Enzyme inhibitors include pancreatic trypsin.

30 Liposomes include water-in-oil-in-water CGF emulsions as well as conventional

liposomes.

The active compound may also be administered parenterally or intraperitoneally. Dispersions can also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations contain a preservative to prevent the growth of microorganisms.

5

The pharmaceutical forms suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. In all cases the form must be sterile and must be fluid to the extent that easy syringability exists. It must be stable 10 under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and vegetable oils. The proper fluidity can be maintained, for 15 example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants.

The prevention of the action of microorganisms can be brought about by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic 20 acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminium monostearate and gelatin.

25 Sterile injectable solutions are prepared by incorporating the active compound in the required amount in the appropriate solvent with various of the other ingredients enumerated above, as required, followed by filtered sterilisation. Generally, dispersions are prepared by incorporating the sterilised active ingredient into a sterile vehicle which contains the basic dispersion medium and the required other ingredients from those 30 enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and the freeze-drying technique which yield a powder of the active ingredient plus any additional desired ingredient from previously sterile-filtered solution thereof.

conventional media or agent is incompatible with the active ingredient, use thereof in the pharmaceutical active substances is well known in the art. Except insofar as any absorption delaying agents and the like. The use of such media and agents for solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and As used herein "pharmaceutically acceptable carrier and/or diluent" includes any and all 30

preparations and formulations. 25
employed. In addition, the active compound may be incorporated into sustained-release form should be pharmaceutically pure and substantially non-toxic in the amounts as cherry or orange flavour. Of course, any material used in preparing any dosage unit as sweetening agent, methyl and propylparabens as preservatives, a dye and flavouring such as shellac, sugar or both. A syrup or elixir may contain the active compound, sucrose as a form of the dosage unit. For instance, tablets, pills, or capsules may be coated with Various other materials may be present as coatings or to otherwise modify the physical 20

contain, in addition to materials of the above type, a liquid carrier. 15
wimergreen, or cherry flavouring. When the dosage unit form is a capsule, it may lactose or saccharin may be added or a flavouring agent such as peppermint, oil of like; a lubricant such as magnesium stearate; and a sweetening agent such as sucrose, phosphatet; a disintegrating agent such as corn starch, potato starch, alginic acid and the The tablets, troches, pills, capsules and the like may also contain the following: a binder 10

therapeutically useful compositions in such that a suitable dosage will be obtained. 5
suspensions, syrups, wafers, and the like. The amount of active compound in such and used in the form of ingestible tablets, buccal tablets, troches, capsules, elixirs, oral therapeutic administration, the active compound may be incorporated with excipients compressed into tablets, or it may be incorporated directly with the food of the diet. For carrier, or it may be enclosed in hard or soft shell gelatin capsules, or it may be orally administered, for example, with an inert diluent or with an assimilable edible When the combination of polypeptides is suitably protected as described above, it may be 55

therapeutic compositions is contemplated. Supplementary active ingredients can also be incorporated into the compositions.

It is especially advantageous to formulate parenteral compositions in dosage unit form for
5 ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the mammalian subjects to be treated; each unit containing a predetermined quantity of active material calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the novel dosage unit forms of the invention are dictated by
10 and directly dependent on (a) the unique characteristics of the active material and the particular therapeutic effect to be achieved, and (b) the limitations inherent in the art of compounding such as active material for the treatment of disease in living subjects having a diseased condition in which bodily health is impaired.

15 The principal active ingredients are compounded for convenient and effective administration in effective amounts with a suitable pharmaceutically acceptable carrier in dosage unit form. In the case of compositions containing supplementary active ingredients, the dosages are determined by reference to the usual dose and manner of administration of the said ingredients.

20

In a further aspect there is provided the combination of the invention as hereinbefore defined for use in the treatment of disease. Consequently there is provided the use of a combination of the invention for the manufacture of a medicament for the treatment of disease associated with aberrant protein/polypeptide structure. The aberrant nature of the
25 protein/polypeptide may be due to misfolding or unfolding which in turn may be due to an anomalous e.g. mutated amino acid sequence. The protein/polypeptide may be destabilised or deposited as plaques e.g. as in Alzheimer's disease. The disease might be caused by a prion. A polypeptide-based medicament of the invention would act to renature or resolubilise aberrant, defective or deposited proteins.

30

The invention is further described below, for the purposes of illustration only, in the following examples.

5 Bacterial and bacteriophage strains. The *E. coli* strains used in this study were: C41(DE3), a mutant of BL21(DE3) capable of expressing toxic genes (Miroux, B. & Walker, J. E. (1996) *J. Mol. Biol.* 260, 289-298); SV2 (B178groEL44), SV3 (B178groEL59) and SV6 (B178groEL673): isogenic strains carrying temperature-sensitive alleles of groEL; SV1(=B178) (Georgopoulos, C., Hendrix, R. W., Casjens, S. (1984) Ph.D. thesis, University of Cambridge, U.K.). Bacteriophage λ b2cl (Tivic, A., Olden, D., Wallington, E. J. & Lund, P. A. (1997) *Gene* 194, 1-8), and TGI R. & Kaiser, A. D. (1973) *J. Mol. Biol.* 76, 45-60), A190 (AgroEL::kanR) [pBAD-EL] (Gibson, T. J. (1984) Ph.D. thesis, University of Cambridge, U.K.). Bacteriophage λ b2cl (Zelitsra-Ryalls, J., Fayet, O., Baird, L. & Georgopoulos, C. (1993) *J. Bacteriol.* 175, 1134-1143) was used according to standard methods (Arber, W., Engquist, L., Hohn, B., Murray, N. E. & Murray, K. (1983) in *Lambda II*, ed. R. W. Hendrix, J. W. r., F. W. Stahl and R. A. Weisberg (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.). 15 T4 (Zelitsra-Ryalls, J., Fayet, O., Baird, L. & Georgopoulos, C. (1993) *J. Bacteriol.* 175, 1134-1143), was used according to standard methods (Karam, J. D. (1994) *Molecular Biology of bacteriophage T4*. (American Society for Microbiology, Washington, DC)); 20 Spring Harbor Laboratory Press, N.Y.)). The schematic organisation of the plasmids used in this study is represented Figure 2. *Gp31* gene was PCR (polymerase chain reaction) amplified using two oligonucleotides 5'-CTTCAGA~~CAT~~ATG TCT GAA GTA CAA CAG CTA CC - 3', and 5' - TAA CGG CCG TTA CTT ATA AAC CGA CCA GAT AGC - 3', producing a 358 bp DNA using pSV25 (van der Vies, S., Gatenby, A. & Georgopoulos, C. (1994) *Nature* 368, 654-656) as template. The DNA sequence of a part of the mobile loop of *Gp31* (residues 25 to 43) was removed by PCR, as described (Hemsley, A., Arneheim, N., Toney, M., D., Cottopass, G. & Galas, D. J. (1989) Nucleic 25 acids 25: 111-118).

Plasmid constructions. Standard molecular biology procedures were used (Sambrook, J.,

plaque formation was assayed at 37°C.

25 Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning. A Laboratory Manual* (Cold Spring Harbor Laboratory Press, N.Y.)). The schematic organisation of the plasmids used in this study is represented Figure 2. *Gp31* gene was PCR (polymerase chain reaction) amplified using two oligonucleotides 5'-CTTCAGA~~CAT~~ATG TCT GAA GTA CAA CAG CTA CC - 3', and 5' - TAA CGG CCG TTA CTT ATA AAC CGA CCA GAT AGC - 3', producing a 358 bp DNA using pSV25 (van der Vies, S., Gatenby, A. & Georgopoulos, C. (1994) *Nature* 368, 654-656) as template. The DNA sequence of a part of the mobile loop of *Gp31* (residues 25 to 43) was removed by PCR, as described (Hemsley, A., Arneheim, N., Toney, M., D., Cottopass, G. & Galas, D. J. (1989) Nucleic acids 25: 111-118).

1. General Experimental Procedures

Examples

Acids Res. 17, 6545-6551), using oligonucleotides 5' – **GGA** GAA GTT CCT GAA CTG – 3' and 5' – **GGA** TCC GGC TTG TGC AGG TTC – 3', creating a unique *BamH* I site (bold characters). *GroEL* gene minichaperone (corresponding to the apical domain of GroEL, residues 191 to 376; (Zahn, R., Buckle, A. M., Perret, S., Johnson, C. M. J., 5 Corrales, F. J., Golbik, R. & Fersht, A. R. (1996) Proc. Natl. Acad. Sci. U.S.A. 93, 15024-15029)) was amplified by PCR using oligonucleotides, containing a *BamH* I site (underlined), 5' – TTC GGA TCC GAA GGT ATG CAG TTC GAC C – 3' and 5' – GTT GGA TCC AAC GCC GCC TGC CAG TTT C – 3' and cloned into the unique *BamH* I site of pRSETA-Gp31Δloop vector, inserting minichaperone GroEL(191-376) in frame 10 into Gp31Δloop sequence. The single ring GroEL_{SR1} mutant contains four amino acid substitutions (R452E, E461A, S463A, and V464A) into the equatorial interface of GroEL, which prevent the formation of double rings (Weissman, J. S., Hohl, C. M., Kovalenko, O., Kashi, Y., Chen, S., Braig, K., Saibil, H. R., Fenton, W. A. & Horwich, A. L. (1995) Cell 83, 577-587). The corresponding mutations were introduced into *groEL* 15 by PCR (Hemsley, A., Arnheim, N., Toney, M. D., Cortopassi, G. & Galas, D. J. (1989) Nucleic Acids Res. 17, 6545-6551) using oligonucleotides 5' – **TGA** GTA CGA TCT GTT CCA GCG GAG CTT CC – 3' and 5' – **ATT** GCG GCG AAG CGC CGG CTG CTG TTG CTA ACA CCG – 3' and pRSETA-*Eag* I GroEL or GroESL vectors (Chatellier, J., Hill, F., Lund, P. & Fersht, A. R. (1998) Proc. Natl. Acad. Sci. U.S.A. 95, 20 9861-9866) as template; silent mutations, in respect to the codon usage in *E. coli*, create a unique *Mfe* I (bold characters) and *Nae* I (underlined). *GroEL*(E191G; *groEL44* allele) gene was PCR amplified from *E. coli* SV2 strain (Zeilstra-Ryalls, J., Fayet, O., Baird, L. & Georgopoulos, C. (1993) J. Bacteriol. 175, 1134-1143) using two oligonucleotides 5' – 25 T AGC TGC CAT ATG GCA GCT AAA GAC GTA AAA TTC GG – 3' and 5' – ATG TAA CGG CCG TTA CAT CAT GCC GCC CAT ACC – 3' producing a 1,659 bp DNA with unique sites for *Nde* I and *Eag* I (underlined). The different genes were subcloned into the unique *Nde* I and *Eag* I unique sites of pACYC184, pJC and pBAD30 (Guzman, L.-M., Belin, D., Carson, M. J. & Beckwith, J. (1995) J. Bacteriol. 177, 4121-4130) vectors (Chatellier, J., Hill, F., Lund, P. & Fersht, A. R. (1998) Proc. Natl. Acad. 30 Sci. U.S.A. 95, 9861-9866). A colony-based PCR procedure was used to identify the positive clones (Chatellier, J., Mazza, A., Brousseau, R. & Vernet, T. (1995) Analyt. Biochem. 229, 282-290). PCR cycle sequencing using fluorescent dideoxy chain terminators (Applied Biosystems) were performed and analysed on an Applied

biosystems 3/3A Automated DNA. All PCR amplified DNA fragments were sequenced after cloning.

7.5 and, dialysed against and stored in 50 mM Tris-HCl, 0.1 mM EDTA, 1 mM β -mercaptoethanol, pH 7.5. Proteins were analysed by electrospray mass spectrometry. Protein concentration was determined by absorbance at 276 nm using the method of Gill & von Hippel (Gill, S. C. & von Hippel, P. H. (1989) *Analyt. Biochem.* 182, 319-326) and confirmed by quantitative amino acid analysis.

Constitutive expression under the control of the tetracycline-resistance gene promoter / operator was obtained using the high copy-number pJC vectors (Chatellier, J., Hill, F., Lund, P. & Fersht, A. R. (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95, 9861-9866). pBAD30 vector allows inducible expression with 0.2-0.5 % arabinose controlled by the P_{BAD} promoter and its regulatory gene, *araC* (Guzman, L.-M., Belin, D., Carson, M. J. & Beckwith, J. (1995) *J. Bacteriol.* 177, 4121-4130). The level of expression of MC₇ was analysed by 15% sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) under non-reducing conditions followed by Western blotting as described (Chatellier, J., Hill, F., Lund, P. & Fersht, A. R. (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95, 9861-9866).

Molecular weight determination by analytical gel filtration chromatography and analytical ultracentrifugation. One hundred μ l aliquots of protein (1 mg. mL^{-1}) were loaded onto a Superdex™ 200 (HR 10/30) column (Pharmacia Biotech.) equilibrated with 50 mM Tris-HCl, 150 mM NaCl, pH 7.5 at 0.5 $\text{mL}\cdot\text{min}^{-1}$ at 20 °C. The column was calibrated using gel filtration standards from Pharmacia Biotech. (thyroglobulin, MW=669 kDa; ferritin, MW=440 kDa; aldolase, MW=158 kDa; ovalbulmin, MW=45 kDa; chymotrypsinogen MW=25 kDa; RNase, MW=13 kDa). Molecular weights were determined by logarithmic interpolation.

Sedimentation analysis was performed in 50 mM Tris-HCl, 2.5 mM DTE (dithio-erythritol), pH 7.2 at 20 °C with protein concentration in the range 45-300 μ M, scanning at 280 nm, with a Beckman XL-A analytical ultracentrifuge, using an An-60Ti rotor. Sedimentation equilibrium experiments were at 10,000 rev. min^{-1} with overspeeding at 15,000 rev. min^{-1} for 6 hours to speed the attainment of equilibrium. Scans were taken at intervals of 24 hours, until successive scans superimposed exactly, when the later scan

tertarmethylbenzidine (TMB, Boehringer Mannheim). Reactions were stopped with 50 μ l conjugate antibodies (Sigma). ELISAs were developed with 3,3',5,5'-coumarinylbenzidine (Sigma) followed by anti-rabbit immunoglobulins horseradish peroxidase antibodies (Sigma) (10 μ g/mL). GroEL molecules were detected with rabbit anti-GroEL GROES proteins (10 μ g/mL). GroEL molecules were detected with rabbit anti-GroEL free peptide solved in 0.1% TFA solution to 1 μ g of proteins prior incubation to coated 30 essentially as above, by adding different concentrations (between 10,000 to 0.1 μ M) of peptide described (Chatellier, J., Buckle, A. M. & Fersht, A. R. (1999) J. Mol. Biol., in press). The inhibition of the binding of MC₇ proteins by the free peptide was analysed by ELISA, 25 C. P., Hendrix, R. W. & Ellis, R. J. (1988) Nature 333, 330-334) was synthesised as Hemmingsen, S. M., Woofford, C., van, d. V. S., Tilly, K., Dennis, D. T., Georgopoulos, A peptide corresponding to the mobile loop of GroES (residues 16 to 32, numbered as in

immunoglobulins horseradish peroxidase conjugated antibodies (Sigma). 20 GroES were detected with rabbit anti-GroES antibodies (Sigma) followed by anti-rabbit MM Tris-HCl, 200 mM KCl, pH 7.4, were bound at 25 °C for 1 hour. Bound GroES were saline: 25 mM NaH₂PO₄, 125 mM NaCl, pH 7.0). GroES, at 10 μ g/mL in 100 μ L of 10 plates were blocked for 1 hour at 25 °C with 2% Marvel in PBS (phosphate buffered 15 4 °C at a concentration of 10 μ g/mL in carbonate buffer (50 mM NaHCO₃, pH 9.6). Proteins were coated onto plastic microtitre plates (Maxisorb, Nunc) overnight at assay). Proteins were coated onto plastic microtitre plates (Maxisorb, Nunc) overnight at 20 GroES binding and competition assays by ELISA (enzyme-linked immunosorbent assay). Proteins were coated onto plastic microtitre plates (Maxisorb, Nunc) overnight at 25 °C

buffer pH 7.8, 2.5 mM DTE (dithioerythrol). 10 GroES were measured on a Jasco J720 spectropolarimeter interfaced with a Neslab PTC-348WI water bath, using a thermostatted cuvette of 0.1 cm path length. Spectra are averages of 10 scans and were recorded with a sampling interval of 0.1 nm. Thermal denaturation was carried out from 5-95 °C at a linear rate of 1 °C/min⁻¹ and monitored at 222 nm. The reversibility was checked after incubation at 95 °C for 20 min and cooling to and equilibrium at 5 °C. The protein concentration was 45 μ M in 10 mL sodium phosphate 15 buffer pH 7.8, 2.5 mM DTE (dithioerythrol).

molecular weight, data were fitted by non-linear regression. 20

was taken as being operational at equilibrium. To evaluate the apparent average

of 1M H₂SO₄ after 10 min and readings taken by subtracting the O.D.650 nm from the O.D.450 nm.

5 **Anti-GroEL antibodies binding by ELISA.** The same amount of proteins (1 µg) were coated as described above. GroEL molecules were detected with either (i) rabbit anti-GroEL horseradish peroxidase conjugate antibodies (9 mg/mL; Sigma) or (ii) rabbit anti-GroEL antibodies (11.5 mg/mL; Sigma) followed by anti-rabbit immunoglobulins horseradish peroxidase conjugate antibodies (Sigma). ELISAs were developed as described above.

10

15 **In vitro refolding experiments.** Refolding assays of pig heart mitochondrial malate dehydrogenase (mtMDH; Boehringer-Mannheim) and aggregation protection were carried out essentially as described (Peres Ben-Zvi, A. P., Chatellier, J., Fersht, A. R. & Goloubinoff, P. (1998) Proc. Natl. Acad. Sci. U.S.A. 95, 15275-15280). The concentrations of MC₇ used were between 8-16 µM (reporter to protomer).

20 **In vivo complementation experiments.** Complementation experiments were performed by transforming electro-competent SV2 or SV6 cells with the pJC series of expression vectors and plating an aliquot of the transformation reactions directly at 43 °C. The percentage of viable cells relative to the growth at 30 °C was determined. A representative number of clones which grew at 43 °C were incubated in absence of any selective markers at permissive temperature. After prolonged growth the loss of the pJC plasmids and the ts phenotype were verified. Each experiment was performed in duplicate. Plasmids carrying no *groE* genes or encoding the GroE proteins were used as 25 negative or positive controls, respectively.

30 P1 transduction (Miller, J. H. (1972) Experiments in Molecular Genetics (Cold Spring Harbor, N.Y.)), using strain AI90 (*ΔgroEL::kan^R*) [pBAD-EL] as donor (Ivic, A., Olden, D., Wallington, E. J. & Lund, P. A. (1997) Gene 194, 1-8), was used to delete the *groEL* gene of TG1 cells transfected by the different pJC vectors. Transductants were selected on LB plates containing 10 µg/mL of kanamycin at 37 °C. Approximately 25 colonies were transferred onto plates containing kanamycin at 50 µg/mL. After incubation for 24 h at 37 °C, colonies that grew were screened by PCR as described.

Gp31loop::GroEL (191-376) (hereafter named MC⁷), where the mobile loop of Gp31 improves their chaperonin-facilitated protein folding, we generated the fusion protein, in an effort to increase the avidity of mimichaperones for substrates, and consequently to

30

this loop by a chosen peptide sequence.
290) projects from each subunit (Figure 1). The basis of the method is the substitution of Chatelier, J., Mazza, A., Brousseau, R. & Vermet, T. (1995) *Analyst Biochem.* 229, 282-90, 361-371). This illustrates that a highly mobile polypeptide loop (residues 25 to 43; 25 ray crystallography (Hunt, J. F., van der Vies, S., Henry, L. & Deisenhofer, J. (1997) *Cell heptameric structure (90 kDa) of which the three-dimensional structure is known from X- product) heptamer. The monomeric protein is 12 kDa, but it spontaneously forms a stable result, the polypeptide is oligomerised. The scaffold is the bacteriophage T4 Gp31 (gene mimichaperone. We describe a scaffold on which any polypeptide may be hung; as a 20 2. Example 1: Gp31 protein as a scaffold for displaying heptameric GroEL*

& Fersht, A. R. (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95, 9861-9866).
76, 45-60) cells was determined essentially as described (Chatelier, J., Hill, F., Lund, P. (Georgopoulos, C., Hendrix, R. W., Casjens, S. R. & Kaiser, A. D. (1973) *J. Mol. Biol.* 286) in T4 (Gibson, T. J. (1984) Ph.D. thesis, University of Cambridge, U.K.) or SV1 vector, Lorist (Gibson, T. J., Rosenthal, A. & Waterston, R. H. (1987) *Gene* 53, 283- Gp31 proteins from pJC vector series on the replication of the bacteriophage λ origin 15 Effect on Lorist's replication of over-expressing of MC⁷. The effect of over-expressing & Fersht, A. R. (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95, 9861-9866).
76, 45-60) cells was determined essentially as described (Chatelier, J., Hill, F., Lund, P. (Georgopoulos, C., Hendrix, R. W., Casjens, S. R. & Kaiser, A. D. (1973) *J. Mol. Biol.* 286) in T4 (Gibson, T. J. (1984) Ph.D. thesis, University of Cambridge, U.K.) or SV1 vector, Lorist (Gibson, T. J., Rosenthal, A. & Waterston, R. H. (1987) *Gene* 53, 283- Gp31 proteins from pJC vector series on the replication of the bacteriophage λ origin 20 effect on Lorist's replication of over-expressing of MC⁷. The effect of over-expressing

10 encoding the GroE proteins were used as negative or positive controls, respectively. Each experiment was performed in triplicate. Plasmids carrying no groE genes or pBAD-EL + pJC vectors] cells on LB plates containing 1% D(+)-glucose or various amounts of arabinose.

5 Transforms were selected at 37 °C on LB supplemented with 50 μg.mL⁻¹ of kanamycin, 120 μg.mL⁻¹ of ampicillin, 25 μg.mL⁻¹ of chloramphenicol and 0.2% L(+)-arabinose. AI90 (AgroEL::kan^R) [pBAD-EL] cells were transformed with the pJC vector series. AI90 (AgroEL::kan^R) [pBAD-EL] cells were transformed with the pJC vector series. 10

was replaced by the sequence of minichaperone GroEL (residues 191 to 376) (Figure 2). MC₇ was cloned downstream of the T7 promoter of pRSETAsht-*Eag* I vector (Chatellier, J., Hill, F., Lund, P. & Fersht, A. R. (1998) Proc. Natl. Acad. Sci. U.S.A. 95, 9861-9866). After sonication, the soluble and insoluble fractions of IPTG-induced transfected 5 C41(DE3) cells (Miroux, B. & Walker, J. E. (1996) J. Mol. Biol. 260, 289-298) were analysed by SDS-PAGE. Most of MC₇ was present in the insoluble fraction. Insoluble material dissolved in 8 M urea was efficiently refolded by dialysis at 4 °C. MC₇ was purified by ion-exchange and gel filtration chromatography. MC₇ was over-expressed in C41(DE3) cells to give 0.25-0.5 g purified protein per L of culture. Purified MC₇ 10 coincided to seven 30.6 kDa subunits of Gp31Δloop::GroEL(191-376) as determined by analytical size exclusion chromatography (Figure 3 a) and analytical ultracentrifugation (Figure 3 b); Gp31Δloop corresponds to a tetra-decamer (14 subunits). The introduction of a foreign polypeptide in the Gp31 scaffold does not prevent its oligomerisation ability. The electronic microscopy studies of MC₇ revealed views that correspond to front views 15 of oligomers with a diameter close to the one of GroEL (J.L. Carrascosa, J.C. & A.R.F., unpublished). The circular dichroism spectrum of MC₇ indicated significant α-helical structure (Figure 4a). The thermal unfolding monitored by far UV-CD was reversible although more than one transition exist (Figure 4b).

20 Bacterial GroES or the human mitochondrial Hsp10 homologous oligomerisable scaffolds have been also successfully used to oligomerise polypeptides displayed in their mobile loops¹.

3. Example 2: Binding to heptameric bacterial co-chaperonin, GroES.

25 The functionality of MC₇ was examined for binding to GroES, since the interaction between GroEL and GroES is known to be less favourable for one monomer than for the heptamer. MC₇ bound specifically to GroES, conversely monomeric minichaperone GroEL(191-376) did not detectably bind the bacterial co-chaperonin (Figure 5a).

30

¹ Please send me details of these experiments.

wild-type GroEL in refolding a non-permissive substrate *in vitro*. (1998) *Molecular Cell* 2, 1-7; this study); remarkably, MC₇ is only 2-fold less active than Valpuesta, J. (1997) *J. Biol. Chem.* 272, 32925-32932; Nilesen, K. L. & Cowan, N. J. GroELSR1 mutant (Lloca, O., Perez-Perez, J., Cartascosa, J., Galan, A., Muñoz, A. & release of GroES to MC₇ (data not shown). Nevertheless, MC₇ is more efficient than of the final yield was observed; indicating the absence of multiple cycles of binding and GroEL (Figure 7c). Although saturating concentration of GroES (4 μ M) does increase about 3- to 5-fold the rates at the beginning of the refolding reaction, a 10-fold decrease 25 mTMH by MC₇ is about 2.5-3 μ M, compared to 6 μ M of enzyme rescued by wild-type mTMH by MC₇ for wild-type GroEL alone (Figure 7b). After 120 min, the yield of refolded 20 mM min⁻¹ refolding denatured mTMH (Figure 7a) with a rate of 0.02 μ M min⁻¹, compared to 0.04 mM min⁻¹ which protect further denatured mTMH from aggregation (Figure 7a) is active in MC₇, but, it is ineffective in enhancing the refolding rate (Figure 7b). Conversely, (Figure 7a) binds denatured mTMH, protecting it from aggregation is monomeric mTMH-chaperone GroEL(191-376) but, it is ineffective in enhancing the refolding rate (Figure 7b). Conversely, Goloubimoff, P. (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95, 15275-15280). Monomeric the co-chaperone GroES (Perez Ben-Zvi, A. P., Chatelier, J., Fersht, A. R. & dehydrogenase (mTMH) refolds in high yield only in the presence of GroEL, ATP, and 15 *In vitro* activity of MC₇. *In vitro*, heat- and diithiothreitol-denatured mitochondrial malate 10 affinity of MC₇ for GroES may be sufficient for multiple binding and release cycles.

Horwitz, A. L. (1996) *Cell* 84, 481-490) is unable to release GroES in the absence of hand, GroELSR1 (Weissman, J. S., Ryer, H. S., Fenlon, W. A., Beecheim, J. M. & cyclizing of GroES on and off GroEL during chaperonin-assisted folding. On the other the formation of the GroEL-GroES complex is low (10⁻⁶ M), which is compatible with μ M compared to 100 μ M for GroEL (Figure 5b). The apparent dissociation constant for synthetic GroES mobile loop peptide did inhibit the binding of MC₇ with an IC₅₀ of 10 loop to displace bound GroES from MC₇, was tested by competition ELISA. The ability of a synthetic peptide corresponding to residues 16 to 32 of GroES mobile 5 μ M of MC₇ to displace bound GroES from MC₇ was tested by competition ELISA. The peptide did inhibit the binding of MC₇ with an IC₅₀ of 10

4. Example 3: *In vivo* complementation of thermosensitive *groEL* mutant alleles at 43 °C.

We sought complementation of two thermosensitive (*ts*) *groEL* mutants of *E. coli* at 43 °C. *E. coli* SV2 has the mutation Glu191→Gly in GroEL corresponding to *groEL44* allele, while SV6 carries the *EL673* allele, which has two mutations, Gly173→Asp and Gly337→Asp. Complementation experiments were performed by transforming the thermosensitive (*ts*) *E. coli* strains SV2 or SV6 with the pJC series of expression vectors vector (Chatellier, J., Hill, F., Lund, P. & Fersht, A. R. (1998) Proc. Natl. Acad. Sci. U.S.A. 95, 9861-9866) and plating an aliquot of the transformation reaction directly at 43 °C. Subsequently, plasmids pJC from a representative number of individual clones growing at 43 °C were lost in the absence of continued chloramphenicol selection. Nearly all ($\geq 95\%$) the cured clones were thermosensitive at 43 °C indicating the absence of recombination events for the reconstitution of wild-type *groEL* gene. The results obtained are qualitatively similar to those previously described (Chatellier, J., Hill, F., Lund, P. & Fersht, A. R. (1998) Proc. Natl. Acad. Sci. U.S.A. 95, 9861-9866). Only minichaperone sht-GroEL(193-335) complements the defect in SV2. The defective *groEL* in SV6 was complemented by expression of minichaperone sht-GroEL(191-345), and less well by sht-GroEL(193-335). Conversely, MC₇ and GroEL_{SR1} complement both temperature-sensitive *E. coli* *groEL44* and *groEL673* alleles at 43 °C (Table 1). Colony-forming units were not observed for either strain at 43 °C with vectors either lacking inserts (pJC_{sht}) or lacking GroEL(191-376) (pJCGp31Δloop).

It has been suggested the higher stability of shortest minichaperone sht-GroEL(193-335) could be responsible for the complementation of *groEL44* mutant allele. To test this eventuality, we purified GroEL(E191G; *groEL44* allele) mutant and compared its thermal stability with the wild-type GroEL. We found no difference in stability between the mutant and the wild-type proteins in presence or absence of ATP. In addition, highly stable functional mutants of GroEL (193-345) do not complement, as the parental minichaperone (Table 1), the defects in SV2 or even SV6. We concluded the thermal stability of minichaperone is not accountable for the complementation of *groEL* defects.

represseed by glucose (Guzman, L.-M., Belin, D., Carson, M. J. & Beckwith, J. (1995) J. 15
 activator depending on the carbon source used. *PBAD* is activated by arabinose but
 promoter and its regulatory gene, *araC*. *AraC* protein acts as either a repressor or an
 plasmid-borne copy of the gene which can be tightly regulated by the arabinose *PBAD*
 chromosomal *groEL* gene has been deleted and *GroEL* is expressed exclusively from a
 complementation of *A190* (*AgroEL::kanR*) [*PBAD-EL*] *E. coli* strain. In this strain, the
 is consistent with the known essential role of *GroEL*. Second, we analysed the
 been deleted, unless intact *GroEL* was expressed from the complementing plasmid. This
 transduction. However, no transductants could be obtained where the *groEL* gene had
groEL gene of *TG1* which had been transformed with the different pJc MC₇ vector by *P1*
groEL gene had been deleted were analysed in two ways. First, we attempted to delete the
 The effects of MC₇ on the growth at 37°C of a strain of *E. coli* in which the chromosomal
 10

10

5

5. Example 4: *In vivo* complementation at 37°C.

<i>GroEL</i> strains	Plasmids pJc	SV2 <i>groEL44</i>	SV6 <i>groEL673</i>
short his tag (sht) (ES, EL)	< 10 ⁻⁴	< 10 ⁻⁴	< 10 ⁻⁴
GP31(I-97)	5 x 10 ⁻³	< 10 ⁻⁴	< 10 ⁻⁴
GP31(I-111)	0.5 x 10 ⁻³	< 10 ⁻⁴	< 10 ⁻⁴
GP31A Loop	< 10 ⁻⁴	< 10 ⁻⁴	< 10 ⁻⁴
GroES(I-548)	1	1	1
GroES-EL			
sht-GroEL(I91-345)	0.01-0.02	0.07-0.09	0.03-0.05
sht-GroEL(I91-376)	< 10 ⁻⁴	< 10 ⁻⁴	< 10 ⁻⁴
sh-GroEL(I93-335)	0.05-0.09	0.15-0.2	0.1
GP31A::GroEL(I91-376)			

strains at 43°C.

Table 1. Relative colony forming ability of transformed *groEL44* or *groEL673* *E. coli*

Bacteriol. 177, 4121-4130). The AI90 [pBAD-EL] cells can not grow on medium supplemented with glucose at 37 °C (Ivic, A., Olden, D., Wallington, E. J. & Lund, P. A. (1997) Gene 194, 1-8). As minichaperones (Chatellier, J., Hill, F., Lund, P. & Fersht, A. R. (1998) Proc. Natl. Acad. Sci. U.S.A. 95, 9861-9866.), MC₇ was unable to suppress this 5 *groEL* growth defect (Table 2). We then determined whether MC₇ could supplement low levels of GroEL from transfected AI90 [pBAD-EL]. At 0.01% arabinose, cells transfected with pJC expressing sht alone, Gp31Δloop or sht-GroEL(191-376), showed little colony forming ability (less than 5%). But those containing pJC MC₇ produced about 30% of the 10 number produced in the presence of 0.2% arabinose. Thus, pJC MC₇, but not pJCGroEL_{SR1}, significantly supplements depleted levels of GroEL., about twice as pJC sht-GroEL(193-335) (Chatellier, J., Hill, F., Lund, P. & Fersht, A. R. (1998) Proc. Natl. Acad. Sci. U.S.A. 95, 9861-9866).

15 **Table 2.** Plating ability of transformed AI90 ($\Delta groEL::kan^R$) [pBADEL] *E. coli* strain at 37 °C in presence of different amount of arabinose.

Plasmids pJC	% L(+)arabinose			
	0.15	0.10	0.01	0.00
short his tag (sht) (<i>ES</i> ⁻ , <i>EL</i> ⁻)	++	+	+/-	-
Gp31Δloop	++	+	+/-	-
GroEL (1-548)	+++	+++	+++	+++
sht-GroEL (191-376)	++	+	+/-	-
Gp31Δ:: GroEL (191-376)	+++	+++	+	-

+++*, growth identical to that in presence of 0.2 % L(+)arabinose (100 %), in terms of both number and size; ++, about 50 % of the colonies relative to that in presence of 0.2*

The *groE* operon was named for its effects on the E protein of λ (Georgopoulos, C., Hendrix, R. W., Casjens, S. R. & Kaiser, A. D. (1973) J. Mol. Biol. 76, 45-60). Although heat induction of the *groE* operon has been shown to decrease burst size of λ bacteriophage in *E. coli* (Wegrzyn, A., Wegrzyn, G. & Taylor, K. (1996) Virology 217, 594-597). In contrast, we showed that the over-expression of GroEL alone, which fold drop in plaques. Over-expression of GroES alone had no effect. Minichaperone GroEL(191-376) had no effect on plaque formation by λ in SV1, but less expression of MC₇ prevents plaque formation by bacteriophage λ in SV1, but less markedly than GroEL (Table 3). It seems that the main effect of GroEL over-expression is mediated through the λ origin, which requires two proteins, O and P. As with GroEL, MC₇ (or GroELSR1) inhibit the replication of the Loris16 plasmid which use the bacteriophage λ origin. The effect on Loris16 shows that the unfoldase activity is also an essential part of GroEL activity *in vivo*. MC₇ and minichaperones possess both, un- and folding, activities. GroEL over-expression gives weak complementation of λ growth in S2 (GroEL44) and S3 (GroEL59; Ser201 \leftrightarrow Phe). MC₇ does not, but GroELSR1 does 20 25 30

37°C (Table 4).

Bacteriophages λ and λ 4 reprogram the chaperonins GroES and GroEL for protein folding during morphogenesis (Zeilstra-Rylls, J., Fayet, O. & Georgopoulos, C. (1991) Annu. Rev. Microbiol. 45, 301-325). Nine *groE* alleles which fail to support λ growth have been sequenced (Zeilstra-Rylls, J., Fayet, O., Baird, L. & Georgopoulos, C. (1993) J. Bacteriol. 175, 1134-1143). We examined the ability of MC₇ over-expressed from the constitutive tet promoter on a high-copy number vector (see Figure 2), to complement three mutant *groE* alleles for plaque formation by λ (b2Cl) at 30 °C (Table 3) and T4 at 10

6. Example 3: Effect on bacteroiophages λ and T4 growth of over-expressing M_{C7}.

% $L(+)$ arabinoose; +, about 30 % of the colonies; +/-, ≤ 5 % of the colonies and size reduced relative to that in presence of 0.2 % $L(+)$ arabinoose; -, no visible colonies.

complement any of the *E. coli* *groEL* mutant strains for bacteriophage λ growth at 30 °C (Table 3).

Bacteriophage T4 (T4D0) also requires a functional *groEL* gene, but encodes a protein

5 Gp31 which can substitute for GroES. The requirement for GroEL can be distinguished genetically from λ 's requirement. Thus only two of the four *groEL* alleles fail to support T4 replication; these are also the two thermosensitive mutations *EL44* and *EL673* (Zeilstra-Ryalls, J., Fayet, O., Baird, L. & Georgopoulos, C. (1993) J. Bacteriol. 175, 1134-1143; Zeilstra-Ryalls, J., Fayet, O. & Georgopoulos, C. (1991) Annu. Rev. Microbiol. 45, 301-325). While over-expression of Gp31 allows T4 growth in all strains (only SV2 and SV6 strains normally do not allow T4 growth), over-expression of Gp31 Δ loop inhibits T4 replication. On the other hand, MC₇ does, as does GroEL_{SR1}, complement *E. coli* *groEL* mutant strains for bacteriophage T4 growth at 30 °C (Table 3).

15 **An unusual basis for thermo-sensitivity in a *groEL* mutation.** Surprisingly, over-expression of GroES demonstrates allele-specific complementation for λ and T4 of GroEL44 (Glu191 \rightarrow Gly) mutant (Tables 3 & 4). The effect is nevertheless incomplete; plaques on SV2 [pJCGroES] are invariably smaller than on SV1, or SV1 [pJCGroES]. The E191G single mutation blocks the assembly of the head structure of bacteriophage λ .

20 A possible molecular basis for this allele-specificity lies in the nature of the *groEL44* mutation. The substitution of Glu191 \rightarrow Gly in the hinge region between the intermediate and apical domains of GroEL presumably increases the flexibility of the hinge, and thereby, modulates a hinged conformational change in GroEL required for proper interaction with GroES. Indeed, the pivoting of the hinge region ensures proper

25 interaction with GroES. For example, the mutant GroEL59 (Ser201 \rightarrow Phe in the same hinge region) in SV3 has low affinity for GroES. Over-expression of GroES will favour the formation of GroES-EL44 complex; we indeed also observed complementation of SV2 for thermosensitivity and bacteriophages growth by over-expressing GroEL44 mutant. Taking advantage of the GroES effect, we observed that GroEL minichaperones

30 and MC₇ all reduce both plaque size and number but, like GroEL, do not completely eliminate them in SV2 [pBADGroES].

GroEL44, purified to homogeneity, is effective in refolding heat- and DTT-denatured mitochondrial maleate dehydrogenase in presence of ATP and saturating concentration of GroES. Surprisingly, GroEL44 is as thermo-stable as the wild-type GroEL, indicating the GroES. Our *in vivo* genetic analysis, the affinity between GroEL44 and GroES is decreased at 37 °C and even more at higher temperature.

Our results suggest that the *groEL44* mutation changes the distribution of GroEL subunits between apical domain-open and closed conformations. To allow GroEL44 to release GroES in the absence of signal transmitted via the binding of ATP to an adjacent ring, we introduced the Glu191→Gly mutation in GroEL44, generating the GroEL44 mutant. GroES is more efficient than MC₇ and even more than GroEL44 *in vitro* and *in vivo*.

Table 3. Growth of bacteriophage λ at 30 °C in transformed wild-type and *groEL* mutant strains.

Plasmids pJC	<i>groEL</i> strains			
	SV1	SV2	SV3	SV6
	(<i>groEL</i> ⁺)	<i>groEL44</i>	<i>GroEL59</i>	<i>groEL673</i>
short his tag (sht) (<i>ES</i> ⁻ , <i>EL</i> ⁻)	+++	-	-	-
GroES (1-97)	+++	+++	-	-
Gp31 (1-111)	+++	-	-	-
Gp31 Δ loop	+++	-	-	-
GroEL (1-548)	-	+	++	+/-
sht-GroEL (191-376)	+++	-	-	-
Gp31 Δ ::GroEL (191- 376)	+	-	-	-

no visible plaques ($<10^4$).
 $+/$, 10²-fold fewer plaques and plaque size reduced relative to wild-type *groEL*⁺ strain; -,
 5 10-fold fewer plaques, or plaque size reduced relative to wild-type *groEL*⁺ strain, or both;
 $+$, 5-fold fewer plaques relative to wild-type *groEL*⁺ strain, or both; +,
 $++$, normal plaque-forming ability relative to wild-type *groEL*⁺ strain, in terms of both
number and size; ++, 5-fold fewer plaques relative to wild-type *groEL*⁺ strain, or both; +,
 $++$, normal plaque-forming ability relative to wild-type *groEL*⁺ strain, in terms of both

376)

Plasmids pJC					short his tag (sht)	(ES ⁻ , EL ⁻)	
SV1	SV2	SV3	SV6	(<i>groEL</i> ⁺)	<i>groEL</i> 44	<i>groEL</i> 59	<i>groEL</i> 673
<i>groEL</i> strains							
-	+++	-	-	+/	+	++	-/+
-	+++	-	+++	+++	+++	+++	-
+++	+++	+++	+++	+++	+++	+++	-
-	+++	+++	+++	+++	+++	+++	-
GP31 (I-III)	GP31 (I-548)	sh-GroEL (I91-376)	GP31Δloop	GP31Δ::GroEL (I91-376)	376)		

Table 4. Growth of bacteriophage T4 at 37 °C in transformed wild-type and *groEL* mutant strains.

Claims

1. A polypeptide monomer capable of oligomerisation, said monomer comprising a polypeptide sequence which potentiates protein folding inserted into the sequence of a subunit of an oligomerisable protein scaffold.
2. A polypeptide monomer according to claim 1, wherein the oligomerisable protein scaffold subunit is selected from the group consisting of bacteriophage T4 Gp31, *Escherichia coli* GroES and homologues thereof of the cpn10 family.
3. A polypeptide monomer according to claim 1 or claim 2, wherein the polypeptide sequence is inserted into the sequence of the oligomerisable protein scaffold subunit such that both the N and C termini of the polypeptide monomer are formed by the sequence of the oligomerisable protein scaffold subunit.
4. A polypeptide monomer according to any preceding claim, wherein the polypeptide sequence is inserted into the oligomerisable protein scaffold subunit by replacing one or more amino acids thereof.
5. A polypeptide monomer according to claim 4, wherein the oligomerisable protein scaffold subunit is bacteriophage T4 Gp31 and the polypeptide sequence is inserted into the oligomerisable protein scaffold subunit by substantially replacing the mobile loop between amino acid positions 27 and 42.
6. A polypeptide monomer according to claim 4, wherein the oligomerisable protein scaffold subunit is *Escherichia coli* GroES and the polypeptide sequence is inserted into the oligomerisable protein scaffold subunit by substantially replacing the mobile loop between amino acid positions 19 and 29.
7. A polypeptide monomer according to claim 4, wherein the oligomerisable protein scaffold subunit is bacteriophage T4 Gp31 and the polypeptide sequence is inserted between positions 59 and 61 of the oligomerisable protein scaffold subunit.

5. A polypeptide monomer according to claim 4, wherein the oligomerisable protein scaffolded between positions 56 and 57 of the oligomerisable protein scaffold is *Escherichia coli* GroES and the polypeptide sequence is inserted between positions 56 and 57 of the oligomerisable protein scaffold subunit.

8. A polypeptide monomer according to claim 4, wherein the oligomerisable protein scaffold subunit is *Escherichia coli* GroES and the polypeptide sequence is inserted between positions 56 and 57 of the oligomerisable protein scaffold subunit.

9. A polypeptide monomer according to claim 4, wherein the oligomerisable protein scaffold subunit is bacteriophage T4 GP31 and polypeptide sequences are inserted at both positions described in claims 5 and 7.

10. A polypeptide monomer according to claim 4, wherein the oligomerisable protein scaffold subunit is *Escherichia coli* GroES and polypeptide sequences are inserted at both positions described in claims 6 and 8.

11. A polypeptide monomer according to claim 2, wherein the polypeptide sequence is displayed at the N or C terminus of the oligomerisable protein scaffold subunit.

12. A polypeptide oligomer comprising two or more polypeptide monomers according to any preceding claim.

13. A polypeptide oligomer according to claim 12, which is a homooligomer.

14. A polypeptide oligomer according to claim 12, which is a heterooligomer.

15. A polypeptide oligomer according to claim 14, wherein complementary protein folding are juxtaposed through the oligomerisation of different polypeptide monomers.

16. A polypeptide oligomer according to any one of claims 12 to 15, wherein the monomers are covalently crosslinked.

17. A polypeptide oligomer according to any one of claims 12 to 16, wherein the protein scaffold is in the form of a ring.

18. A polypeptide oligomer according to claim 17, wherein the ring is a heptameric ring.
- 5 19. A polypeptide monomer or oligomer according to any preceding claim, wherein the polypeptide sequence is selected from the group consisting of a minichaperone, a protease prosequence and a foldase.
- 10 20. A polypeptide oligomer or monomer according to claim 19, wherein the foldase is selected from the group consisting of a thiol/disulphide oxidoreductase and a peptidyl prolyl isomerase.
- 15 21. A method for promoting the folding of a polypeptide comprising contacting the polypeptide with a polypeptide oligomer or monomer according to claim 19 or claim 20.
22. A method according to claim 21, wherein the polypeptide is an unfolded or misfolded polypeptide.
- 20 23. A method according to claim 21 or 22, wherein the polypeptide comprises a disulphide.
24. A method according to any one of claims 21 to 23, wherein the foldase is selected from the group consisting of thiol/disulphide oxidoreductases and peptidyl-prolyl isomerasess.
- 25 26. A method according to claim 24, wherein the thiol/disulphide oxidoreductase is selected from the group consisting of *E. coli* DsbA and mammalian PDI, or a derivative thereof.
- 30 26. A method according to claim 24, wherein the peptidyl prolyl isomerase is selected from the group consisting of cyclophilin, parbulen, SurA and FK506 binding proteins.

27. A method according to any one of claims 21 to 26, comprising contacting the polypeptide with a polypeptide oligomer according to any one of claims 12 to 20 and a foldase and a non-oligomerised foldase.

28. A method according to any one of claims 21 to 27, wherein the polypeptide oligomer according to any one of claims 21 to 26 and/or the foldase is immobilised onto a solid phase support.

10 29. A method according to claim 28 wherein the solid phase support is agarose.

30. A solid phase support having immobilised thereon a polypeptide oligomer according to any one of claims 12 to 20 and/or a foldase.

15 31. A column packed at least in part with a solid phase support according to claim 30.

32. Use of a polypeptide according to any one of claims 12 to 20, optionally in combination with a foldase, for promoting the folding of a polypeptide.

20 33. Use according to claim 32 wherein the polypeptide according to any one of claims 12 to 20 and/or the foldase is immobilised on a solid phase support.

34. A composition comprising a combination of a polypeptide oligomer according to any one of claims 12 to 20 and a foldase.

Abstract

The invention relates to polypeptide monomer capable of oligomerisation, said monomer comprising a polypeptide which potentiates protein folding inserted into the sequence of a subunit of an oligomerisable protein scaffold.

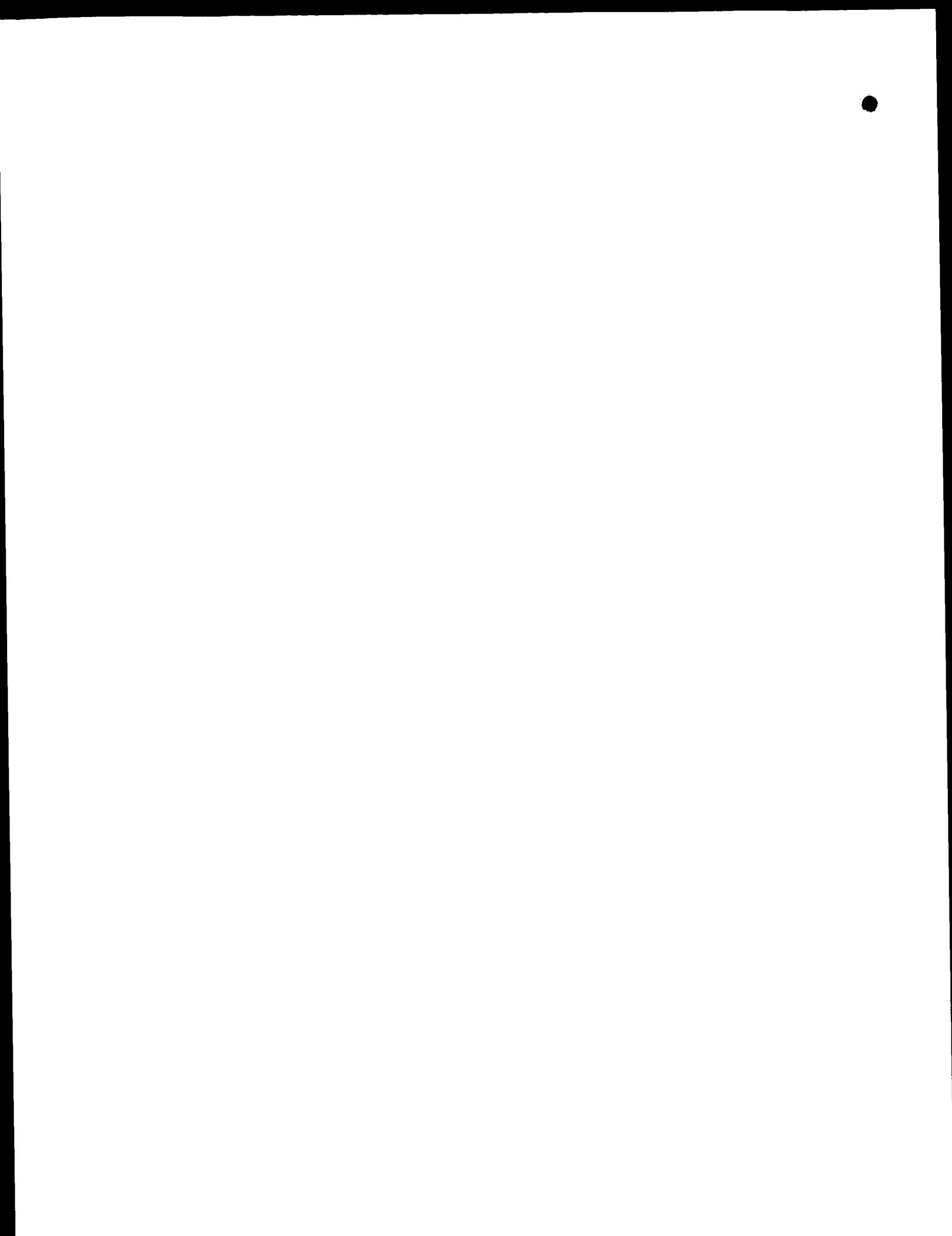
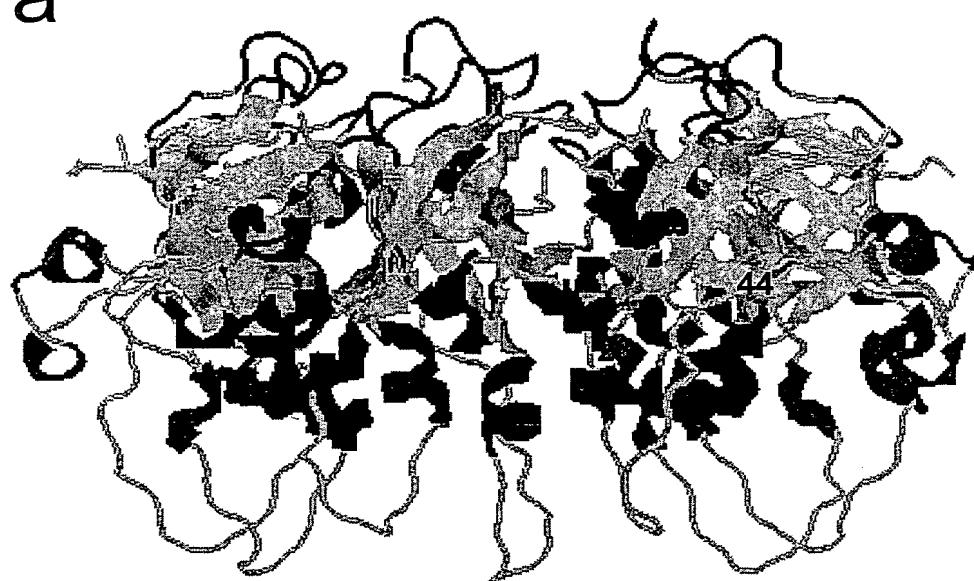


Figure 1.

a



b



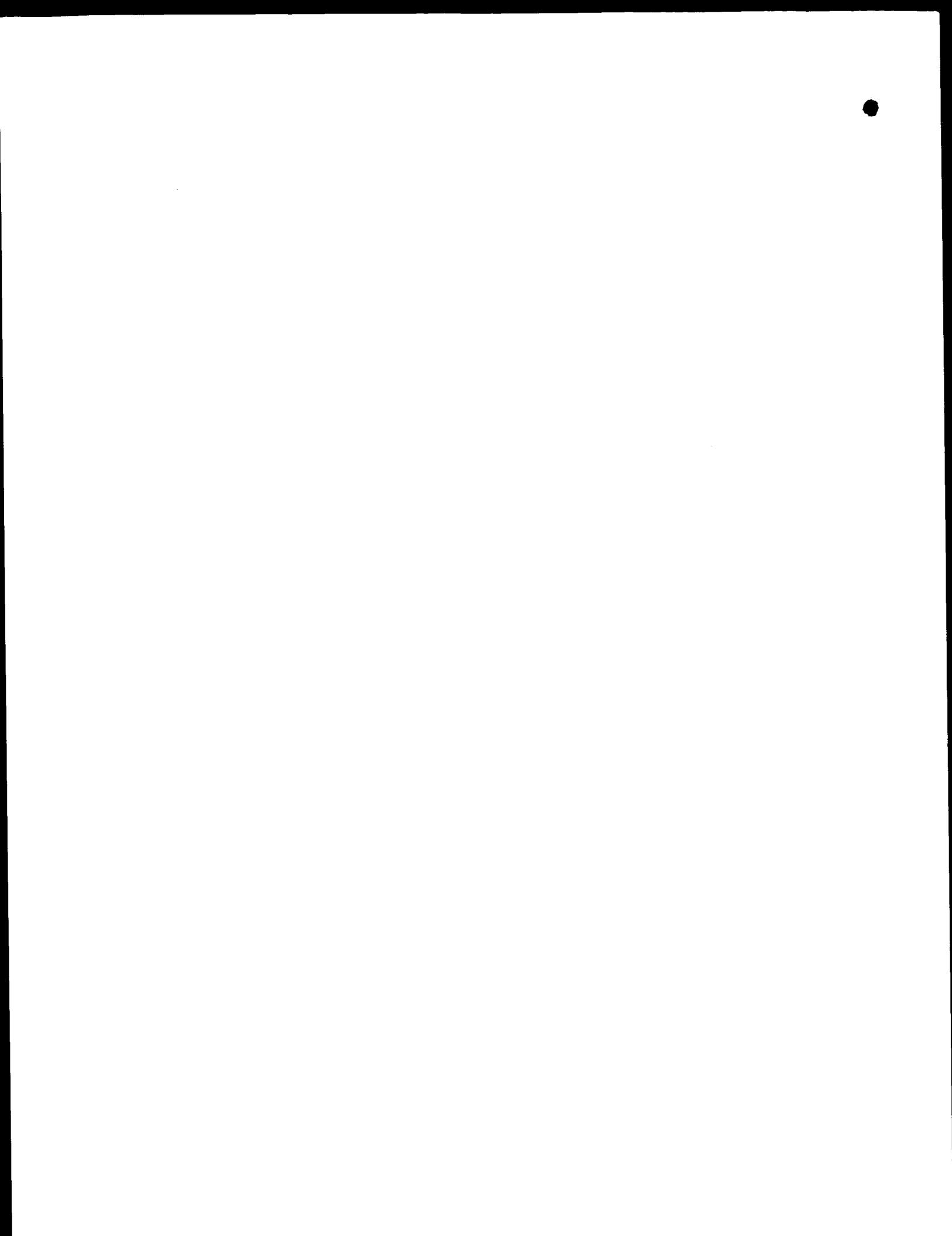
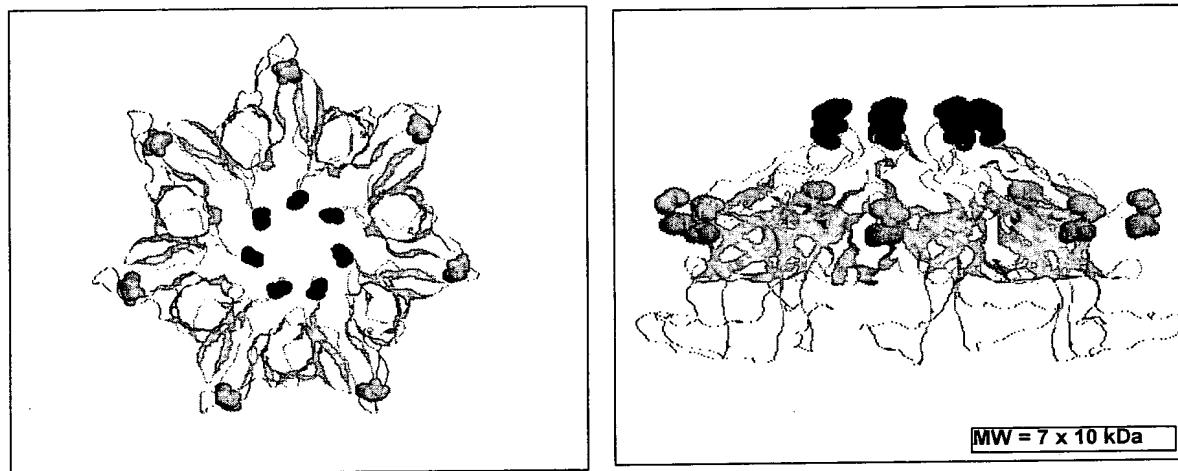


Figure 2



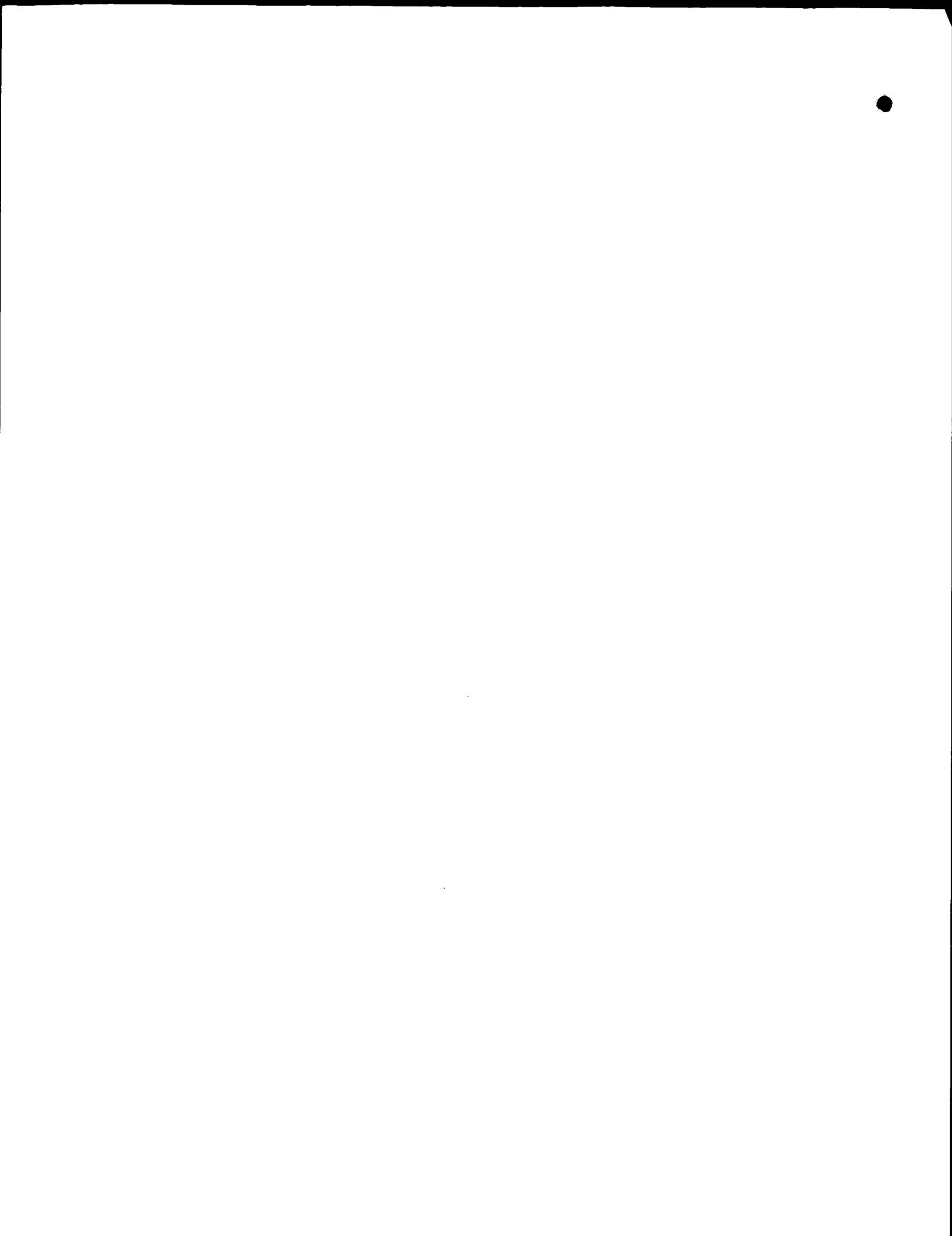
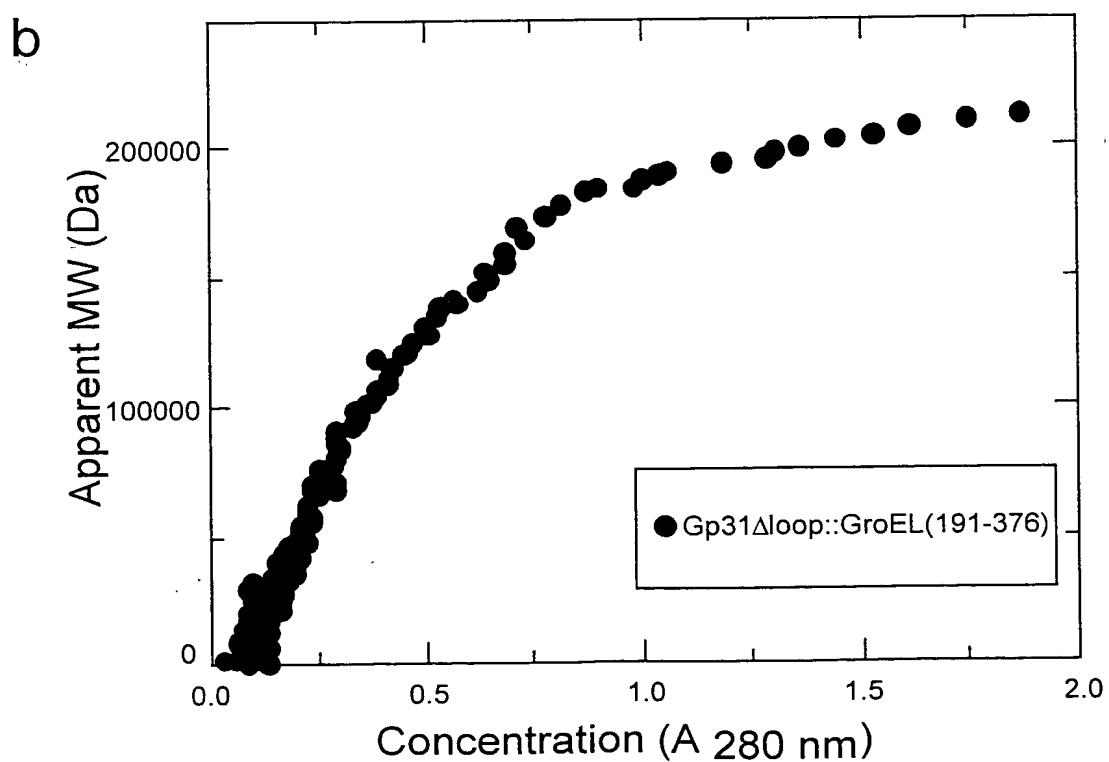
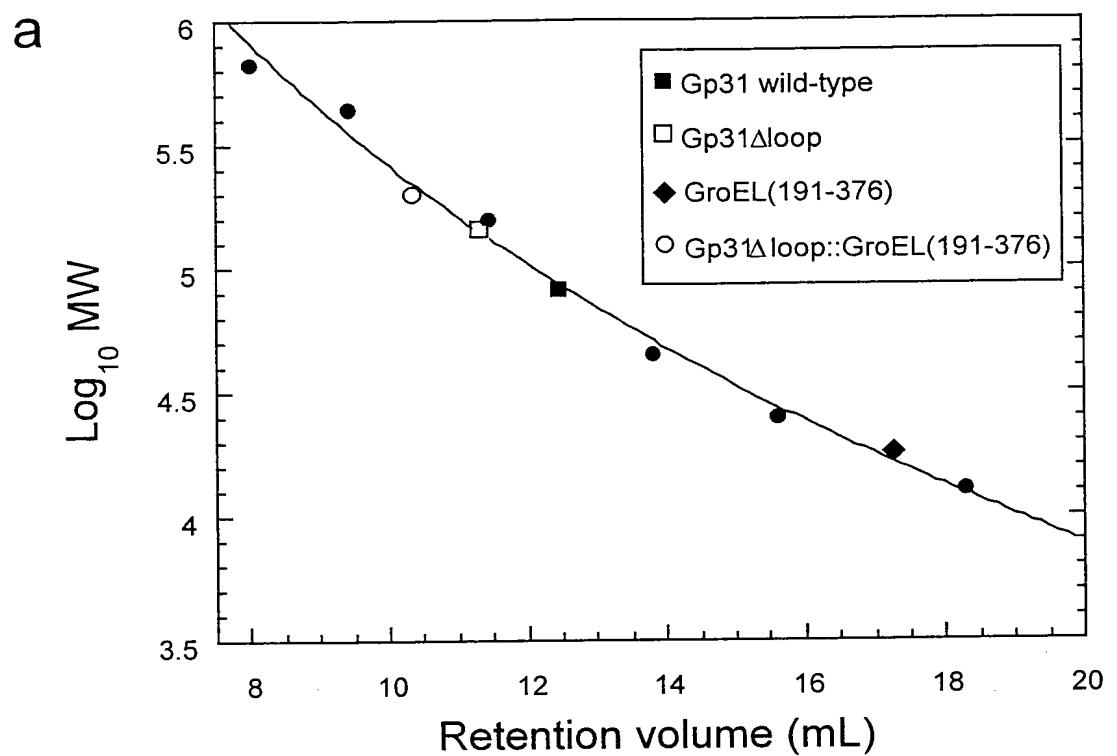


Figure 3.



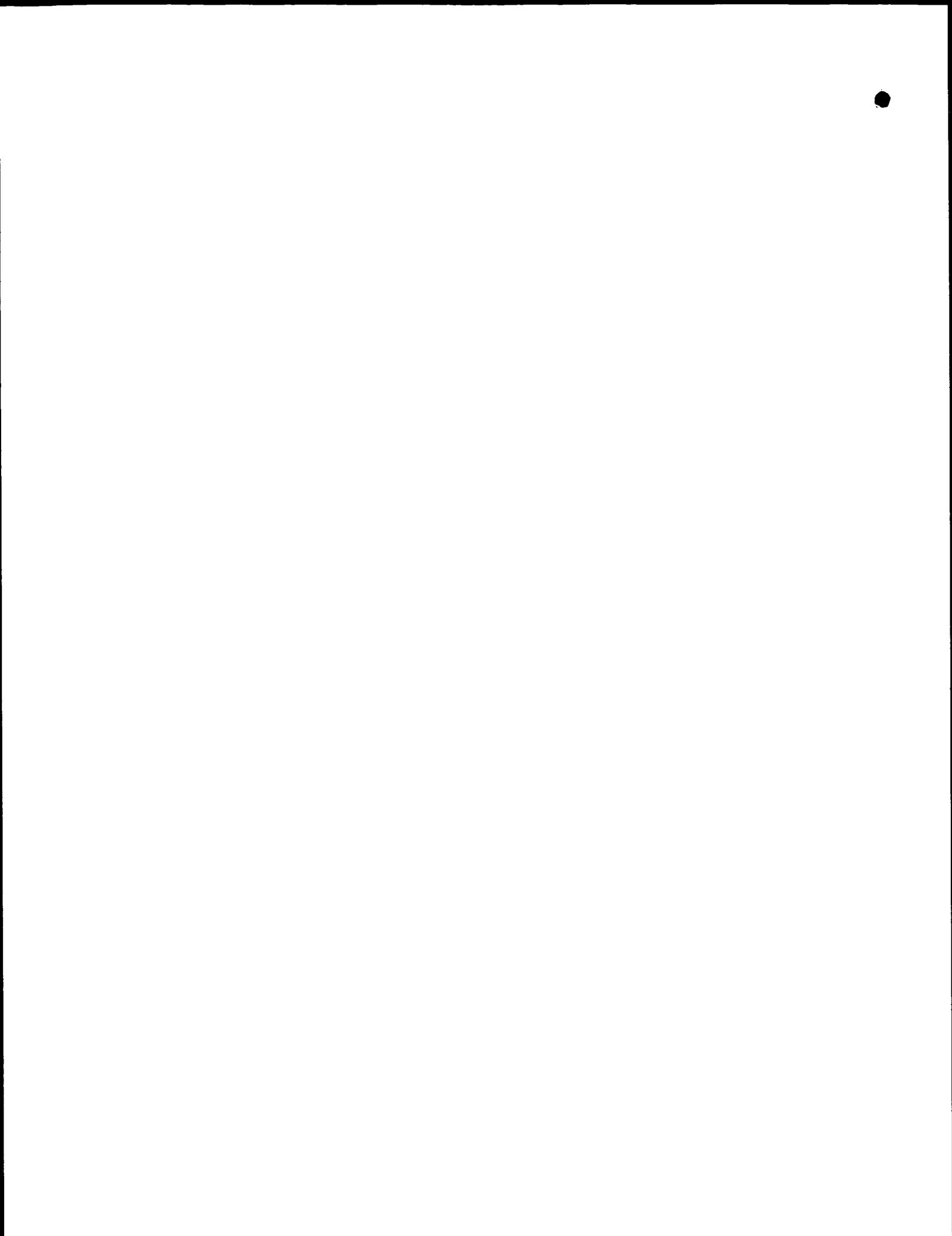
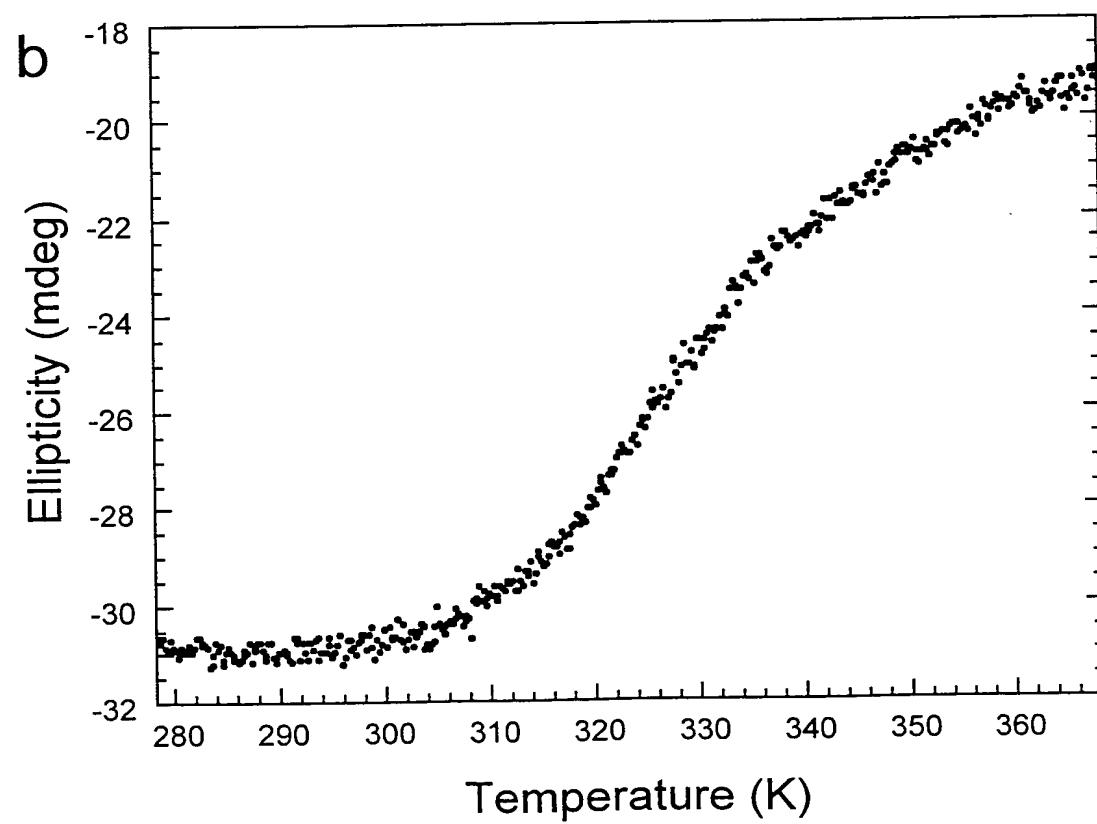
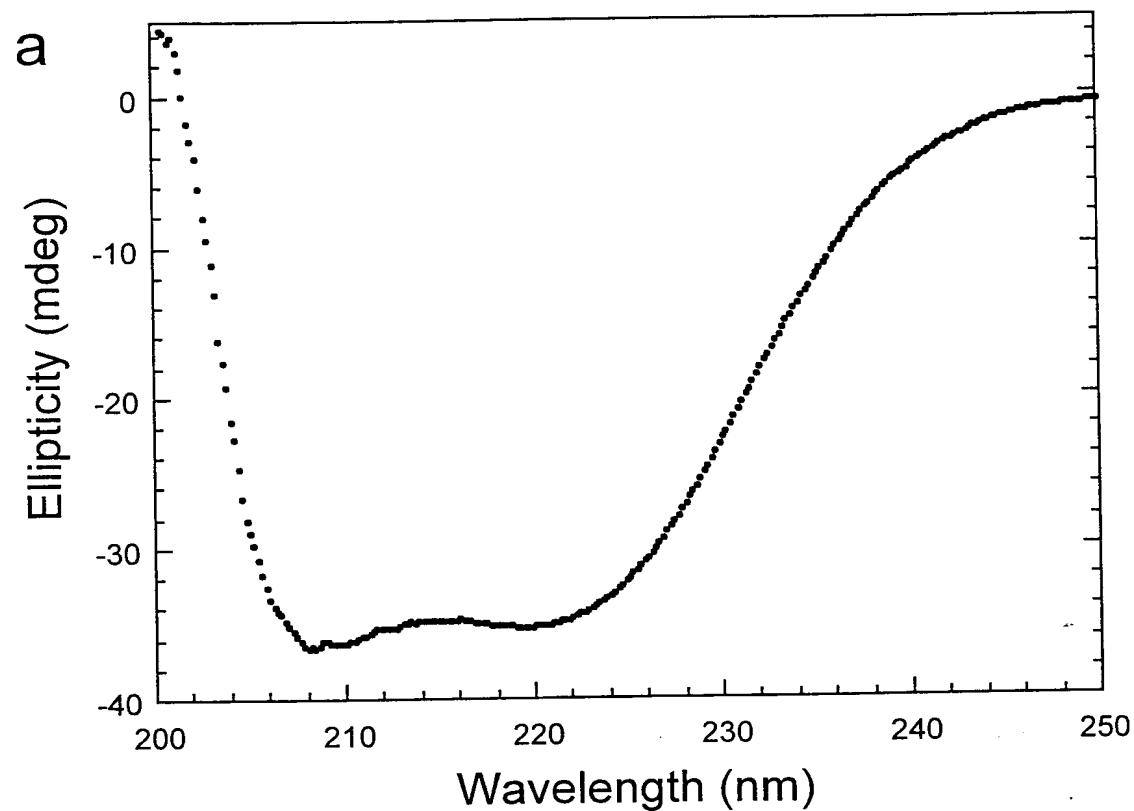


Figure 4.



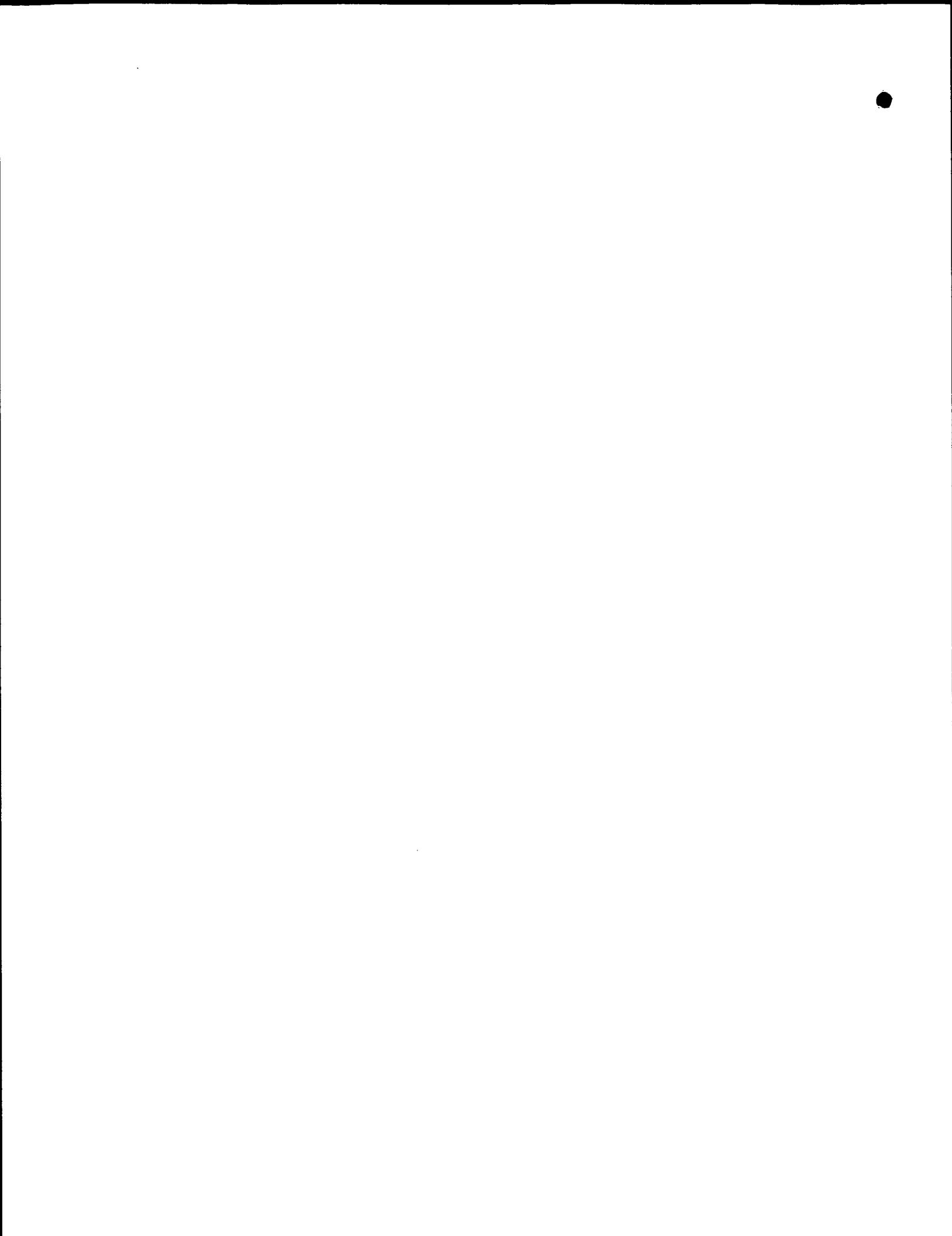


Figure 5.

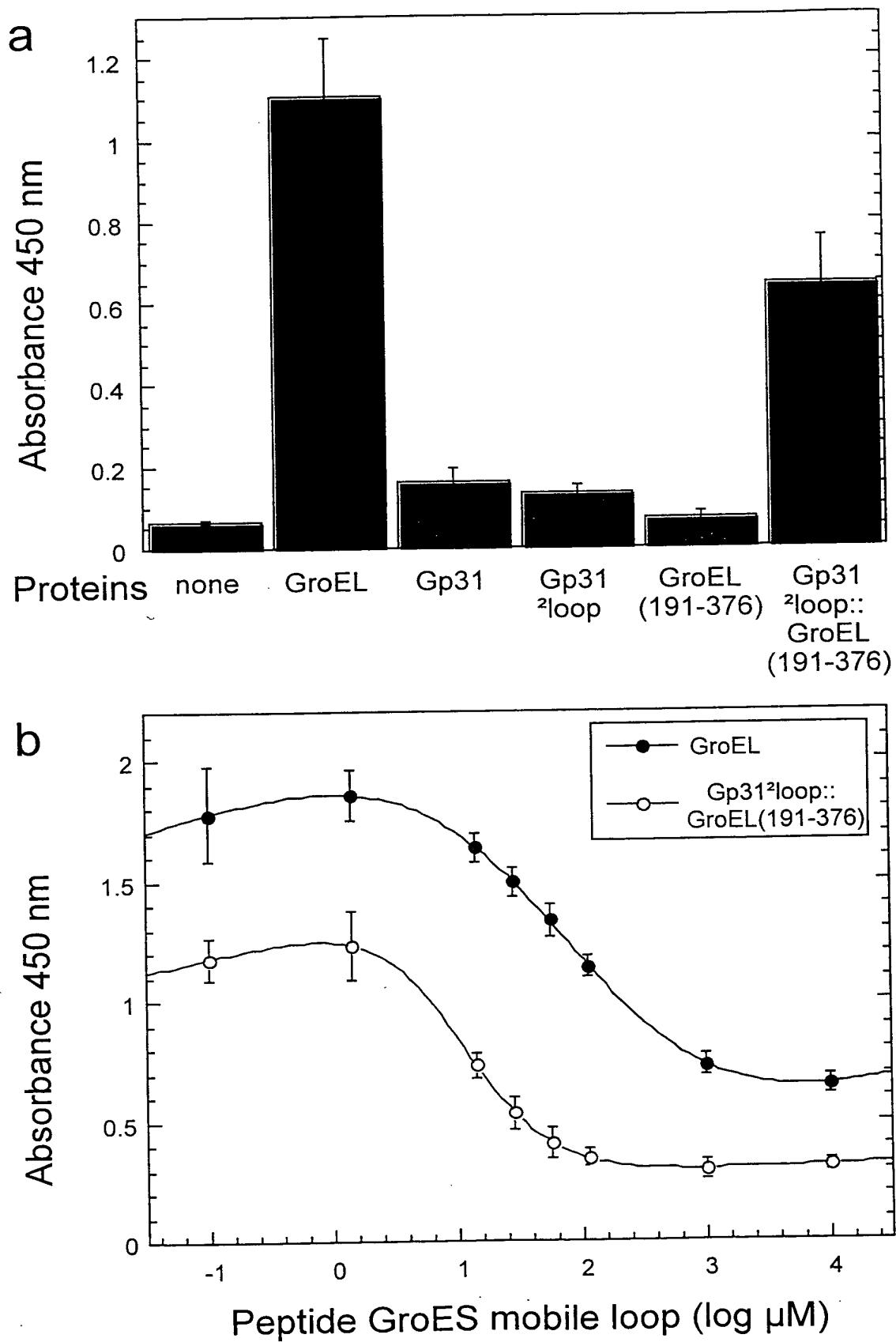
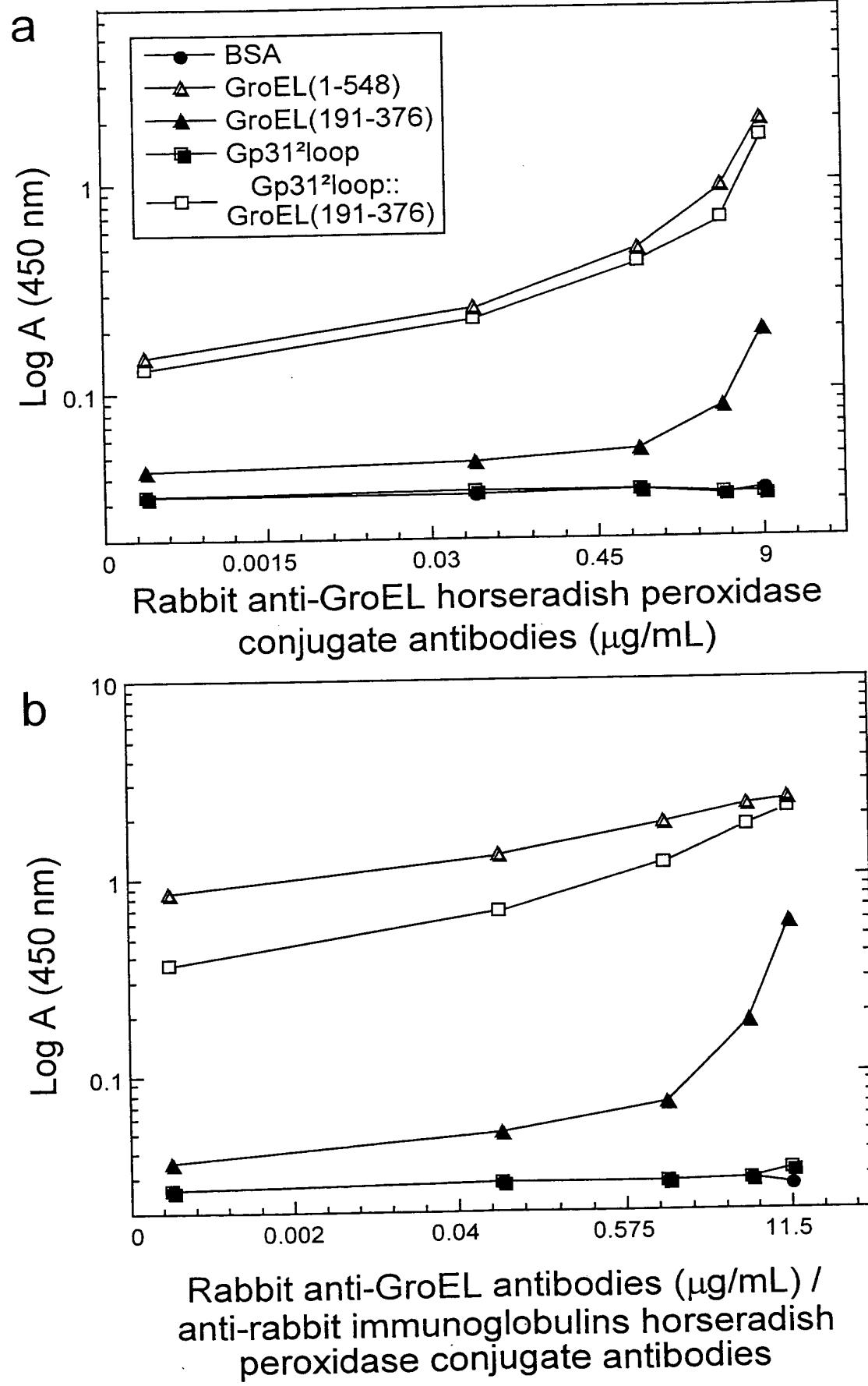




Figure 6.



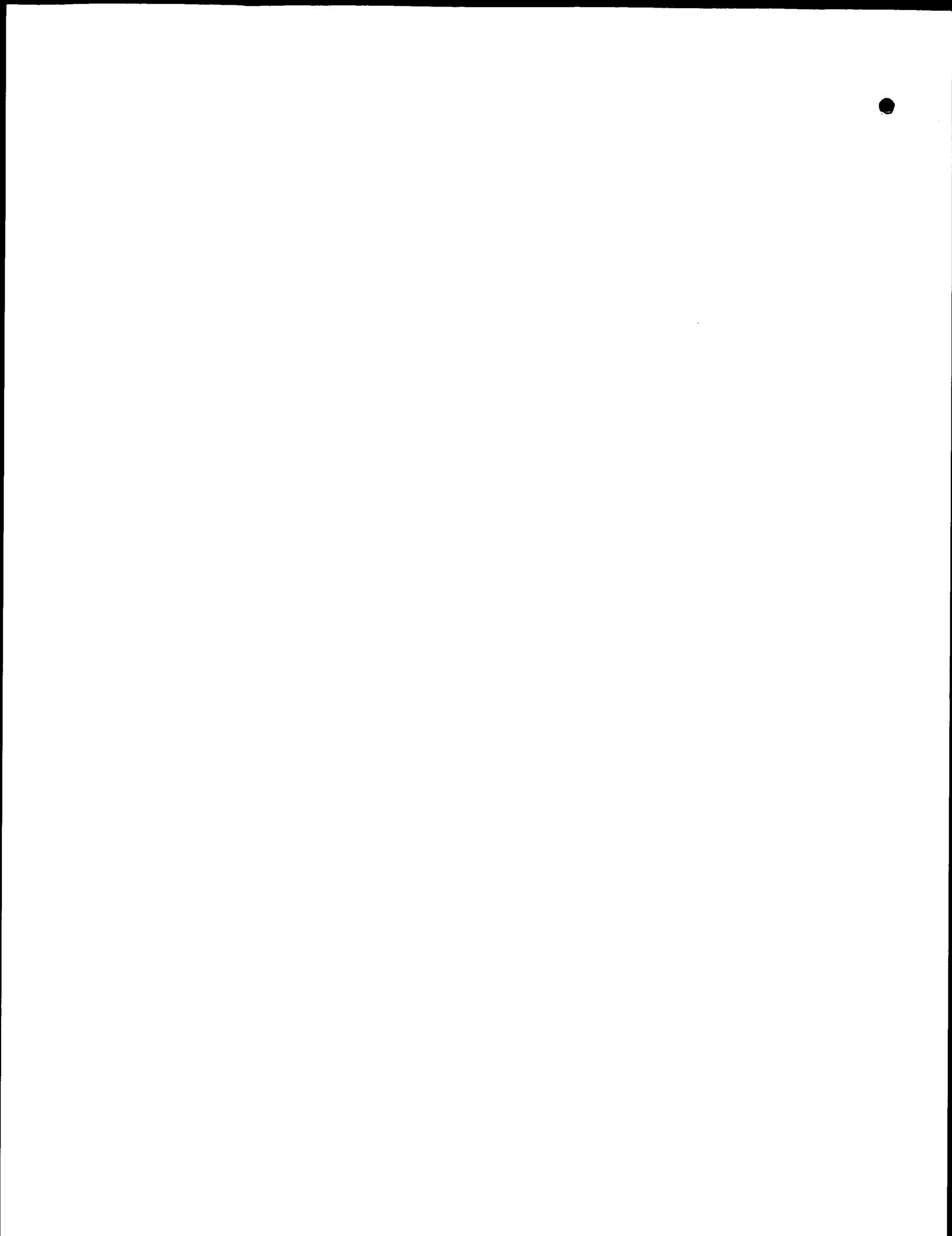
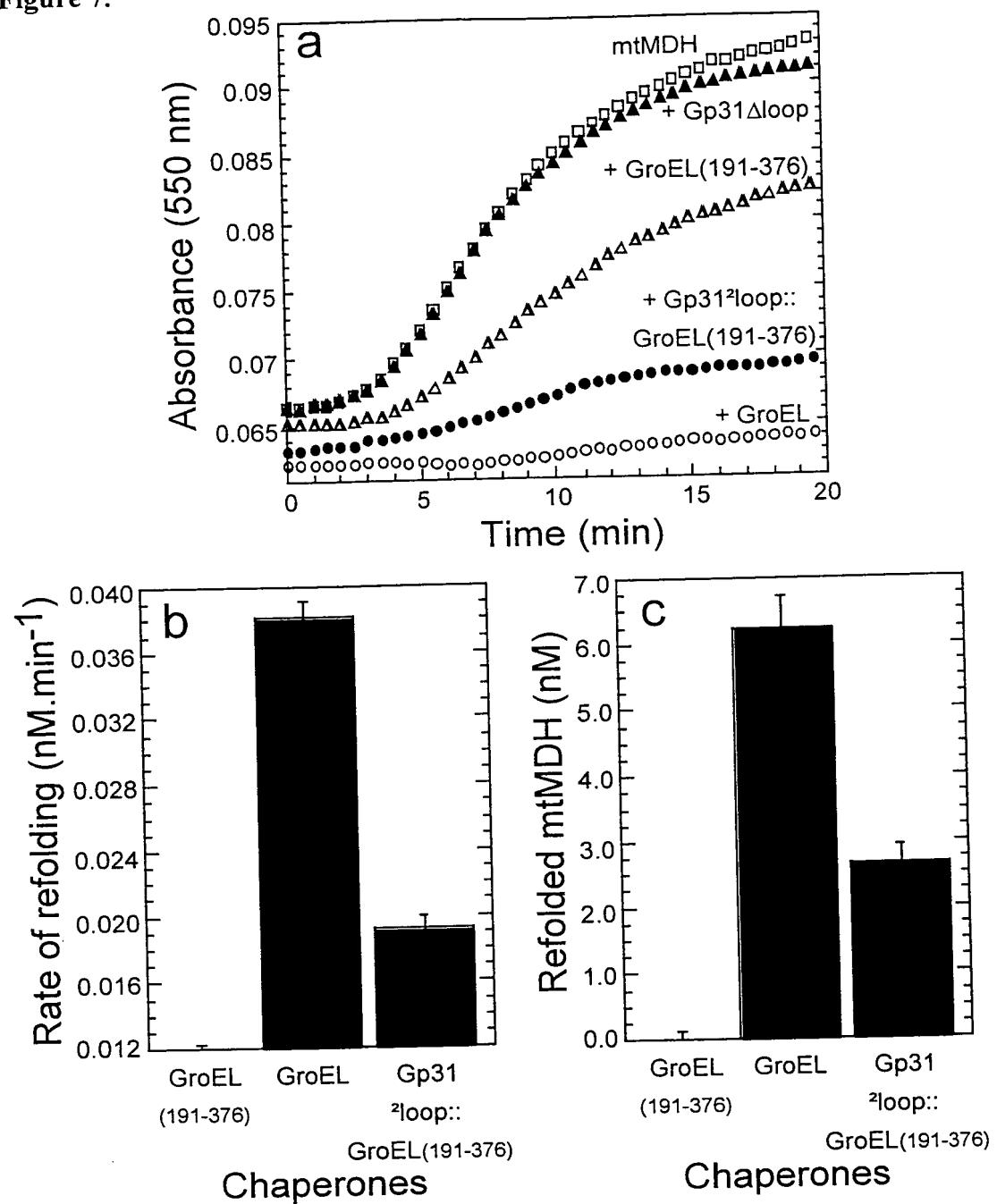


Figure 7.



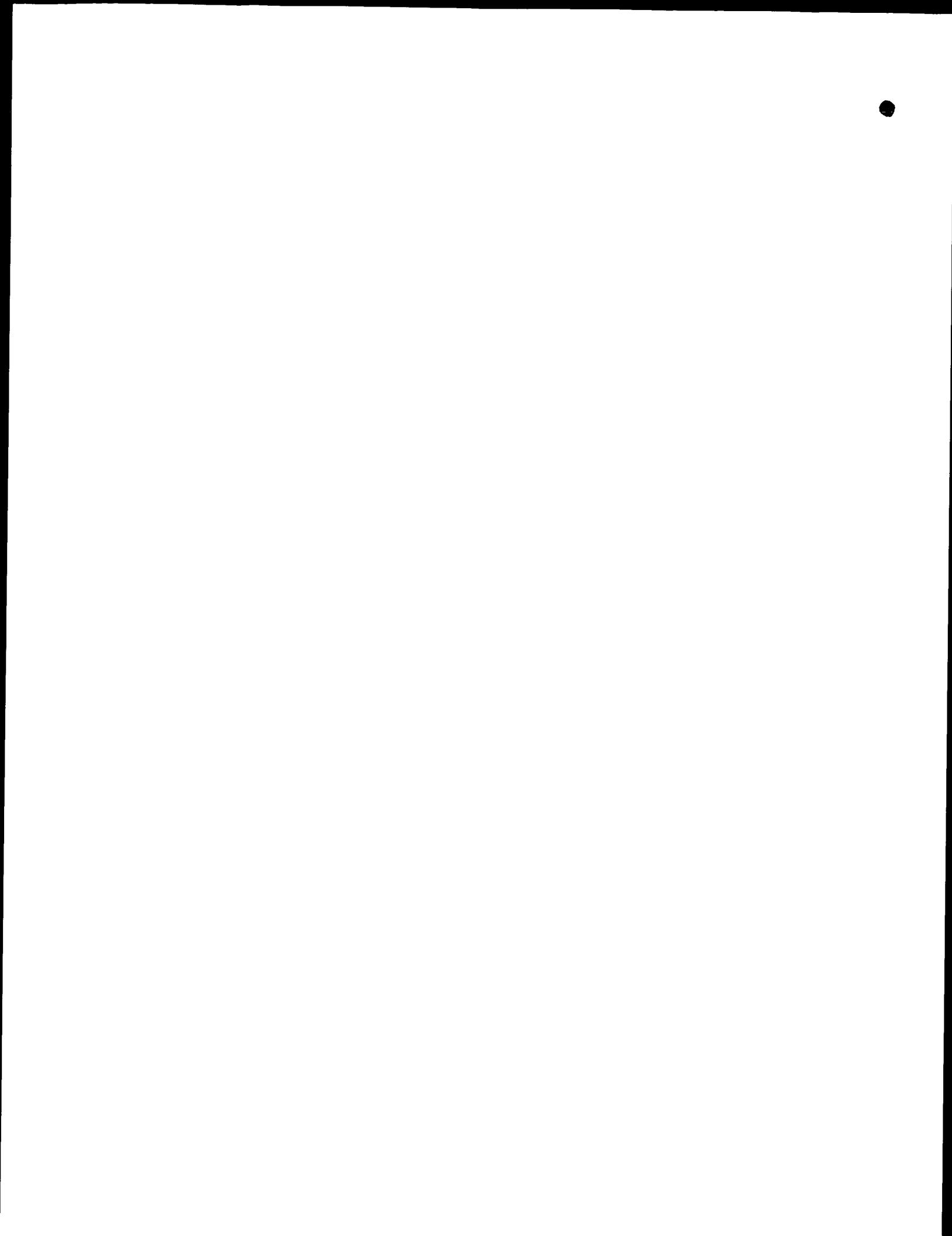
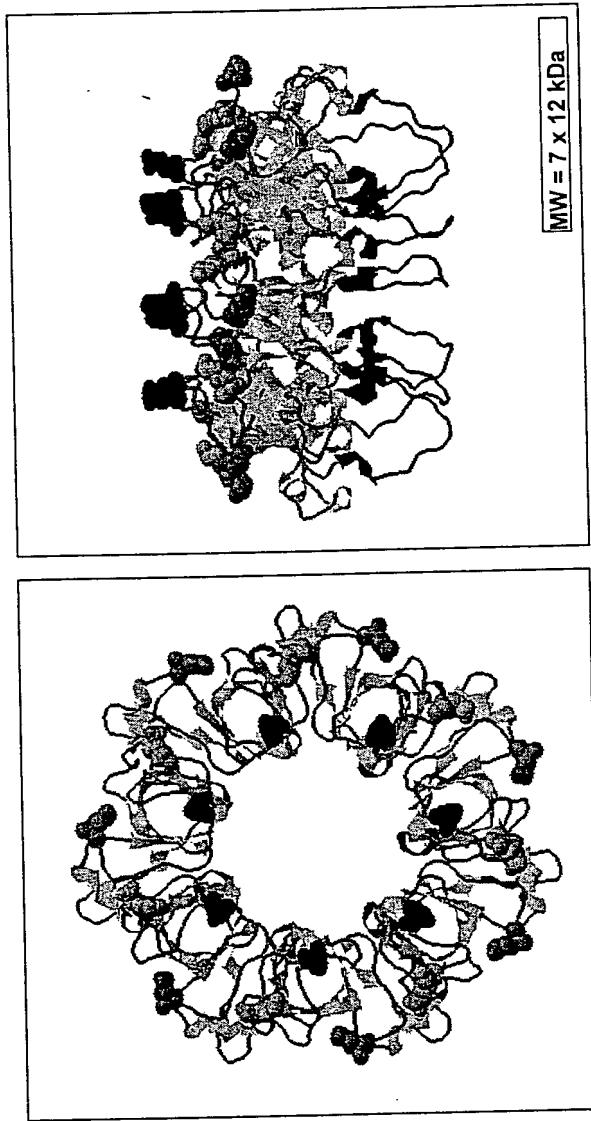


Figure 8



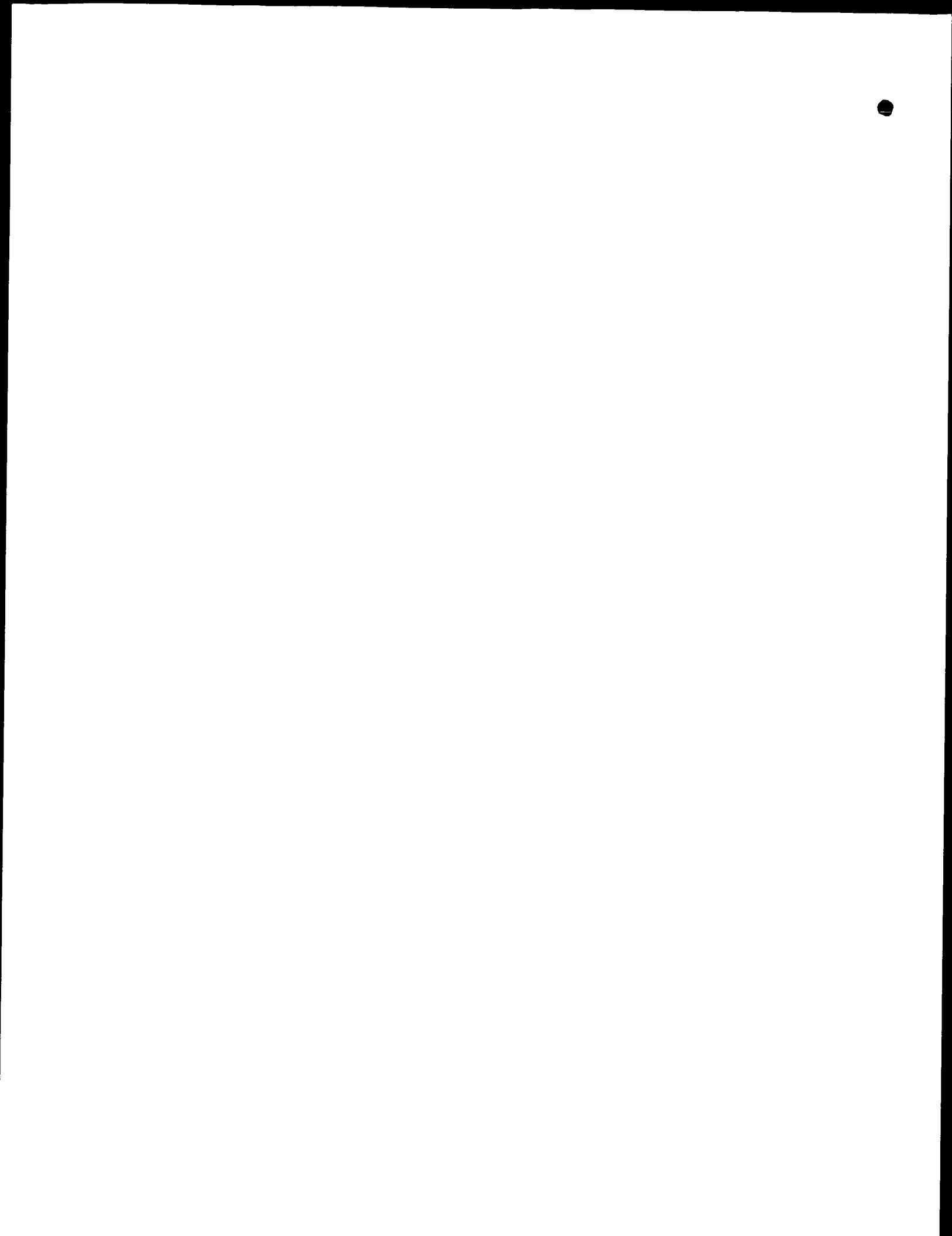
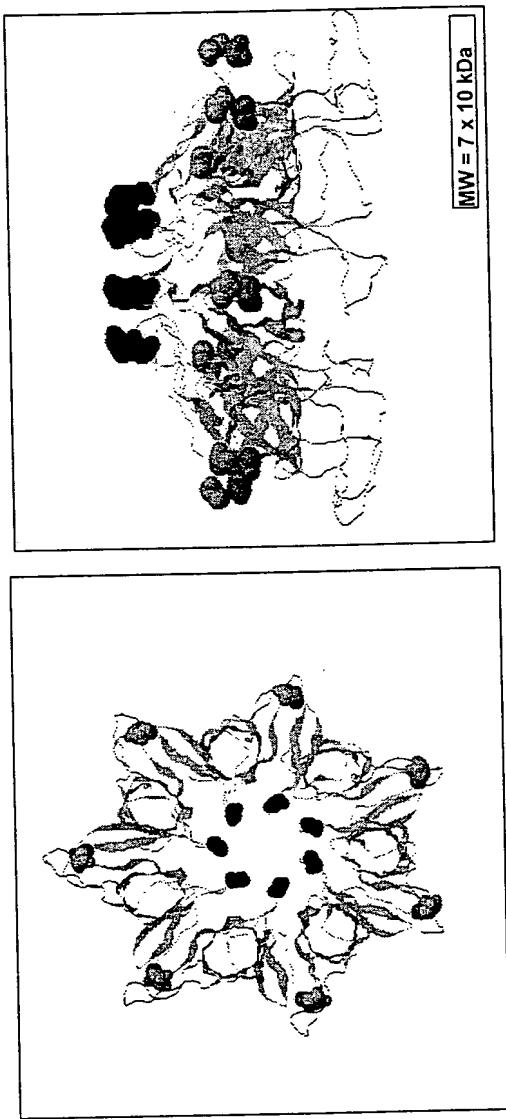
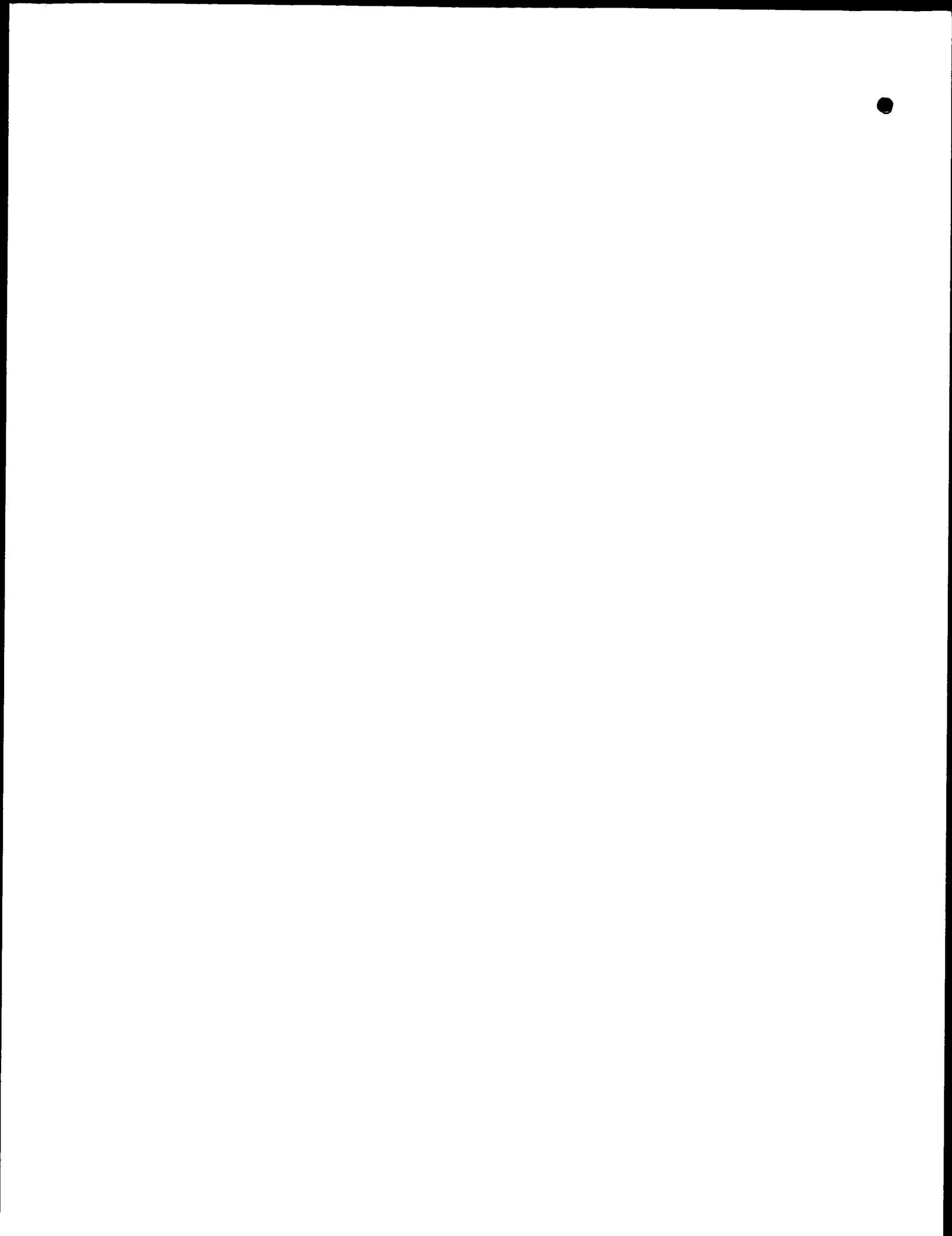
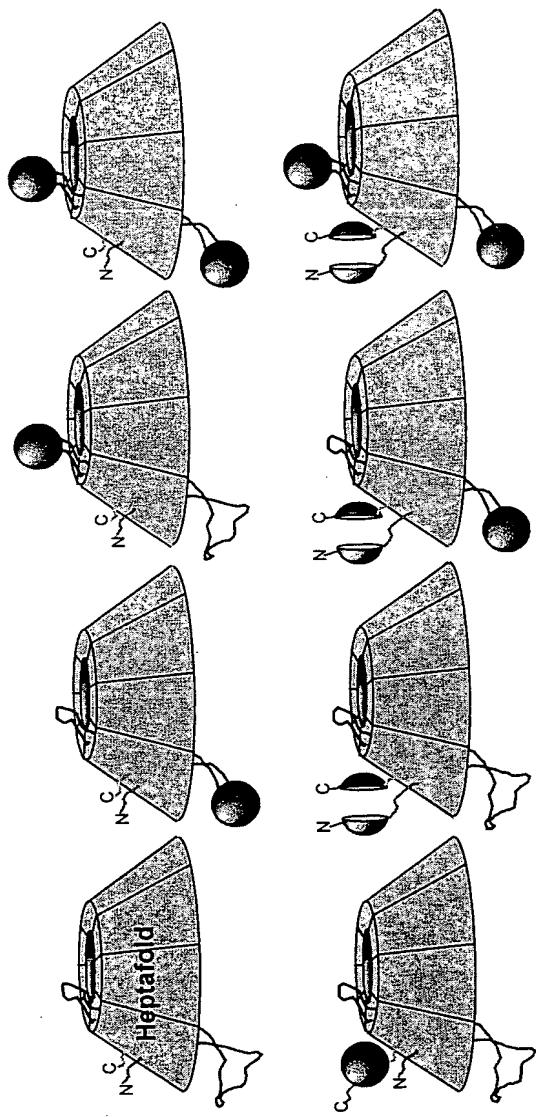


Figure 9





Topologies of Heptafold



Total number of combinations = $2 \times (2 \times 2 \times 2) = 5 \cdot 4 \cdot 10^8$

Figure 10

